AWARD NUMBER: W81XWH-19-1-0621
BM180109

TITLE: Unraveling the Role of DDX41 in Hematopoiesis and MDS

PRINCIPAL INVESTIGATOR: Teresa Bowman

CONTRACTING ORGANIZATION: Albert Einstein College of Medicine

REPORT DATE: October 2020

TYPE OF REPORT: ANNUAL

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2020	ANNUAL	09/01/2019–08/31/2020

4. TITLE AND SUBTITLE

Unraveling the Role of DDX41 in Hematopoiesis and MDS

5a. CONTRACT NUMBER
W81XWH–19–1–0621

5b. GRANT NUMBER
BM180109

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Teresa Bowman

E-Mail: Teresa.bowman@einsteinmed.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Albert Einstein College of Medicine

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

DOD/BMFRP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Myelodysplastic syndromes (MDS) are bone marrow failure disorders arising from defects in blood-forming stem cells resulting in stem cell expansion and decreased production of mature blood cells. MDS in adults is largely thought to be acquired, but the recent identification of a familial form of adult-onset MDS linked to germline mutations in the *DEAD-box Helicase 41* gene (*DDX41*) challenges this dogma. Although this form of MDS is rare, lessons learned from familial forms of human diseases often provide information that is more broadly applicable to acquired forms of a disease that arise from sporadic and/or complex genetic changes. Germline mutations in *DDX41* disrupt its function such that insufficiency of normal levels of DDX41 is thought to contribute to MDS formation. DDX41 plays several roles in cellular homeostasis, thus it is unclear which aspects are aberrant when DDX41 is mutated and how dysfunction of these processes contribute to MDS. To address this question, we have established a zebrafish model of *ddx41* deficiency, which is one of the first animal models for studying the *in vivo* requirements of Ddx41 in hematopoiesis. Human and zebrafish DDX41 are highly conserved, thus we anticipate that our studies in zebrafish will be informative for understanding the function of human DDX41 in hematopoiesis. Zebrafish is an easily manipulated vertebrate model system that has been used extensively to study the basis of both benign and malignant hematopoiesis including MDS. We aim to utilize the advantages of the zebrafish to uncover how Ddx41 regulates hematopoiesis and identify novel therapeutic targets for the treatment of MDS.

15. SUBJECT TERMS

Myelodysplastic syndrome, DDX41, hematopoiesis, inflammation, DNA damage

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	88	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

	Page
1. Introduction	4
2. Keywords	4
3. Accomplishments	4-7
4. Impact	7-8
5. Changes/Problems	8
6. Products	8
7. Participants & Other Collaborating Organizations	8-9
8. Special Reporting Requirements	9
9. Appendices	9-88

1. Introduction

Myelodysplastic syndromes (MDS) are a spectrum of disorders arising from hematopoietic stem and progenitor cell (HSPC) dysfunction resulting in ineffective hematopoiesis and cytopenias. MDS in adults is largely thought to be an acquired disorder, but the recent identification of an inherited form of adult-onset MDS linked to germline mutations in the *DEAD-box Helicase 41* gene (*DDX41*) challenges this dogma. Germline mutations in *DDX41* primarily occur in the N-terminal domain leading to an early stop codon, suggesting that diminished function of DDX41 contributes to MDS initiation. DDX41 is a nucleic acid-activated ATPase linked to immunity, splicing, and R-loop regulation. All three of these processes have been implicated in MDS, thus the underlying mechanism for how DDX41 mutations contribute to hematologic malfunctioning is unclear. Our central *premise* is that the varied functions of DDX41 are connected and necessary for hematopoietic homeostasis. To discover how DDX41 insufficiency could lead to MDS, we established a zebrafish model of *ddx41* deficiency, which is one of the first animal models for studying the *in vivo* requirements of Ddx41 in hematopoiesis. Human and zebrafish DDX41 are highly conserved sharing >85% amino acid identity, thus we anticipate that our studies in zebrafish will be informative for understanding the function of human DDX41 in hematopoiesis. Zebrafish is a genetically malleable vertebrate model system that has been used extensively to study the basis of both benign and malignant hematopoiesis including MDS. Zebrafish mutants for *ddx41* display HSPC expansion as well as neutropenia and anemia, phenotypes reminiscent of human MDS. We aim to utilize the advantages of the zebrafish to uncover how Ddx41 regulates hematopoiesis and identify novel therapeutic targets for the treatment of MDS.

2. Keywords

Myelodysplastic syndrome, DDX41, hematopoiesis, inflammation, DNA damage

3. Accomplishments

Major goals:

Goals as outlined in the approved Statement of Work:

	Timeline Months	**Site 1**
Task: Seek regulatory approval for the use of animals	1-3	Local IACUC and DoD ACURO
Aim 1: To determine R-loop involvement in Ddx41 regulation of hematopoiesis formation		
1.1 To determine if DDX41 regulates R-loop levels.	1-6	Bowman lab
1.2 To determine if R-loops contribute to hematopoietic defects in *ddx41* mutants.	1-18	Bowman lab
1.3 To determine how R-loops regulate hematopoietic gene expression.	1-24	Bowman lab
Milestone Achieved: Delineate how DDX41-mediated R-loop regulation contributes to blood defects.	24	
Aim 2: To Determine R-loop-triggered DNA damage responses during development.		
2.1 To determine if inflammatory signaling is perturbed in *ddx41* mutants.	1-12	Bowman lab
2.2 To determine if STING-mediated inflammatory signaling plays a role in hematopoietic defects in *ddx41* mutants.	6-24	Bowman lab
2.3 To determine if elevated levels of R-loops alter inflammatory signaling in *ddx41* mutants.	12-18	Bowman lab
Milestone Achieved: Define how DDX41 insufficiency alters the inflammatory milieu and the impact of this on hematopoiesis	24	

Progress towards year 1 goals:

Task: Seek regulatory approval for the use of animals: The project was approved by both the IACUC at Albert Einstein College of Medicine (12/18/19) and the DOD ACURO (2/24/20).

Aim 1.1: To determine if DDX41 regulates R-loop levels. To decipher if Ddx41 plays a role in R-loop regulation *in vivo*, we examined R-loop levels in zebrafish *ddx41* mutants by using the monoclonal antibody S9.6 that specifically recognizes RNA:DNA hybrid structures. RNA:DNA hybrid immunofluorescence signal in embryonic

cells was significantly higher in *ddx41* mutants compared to siblings. To validate that the structures examined are indeed RNA:DNA hybrids, we demonstrated that the S9.6 levels were sensitive to RNASEH1, an enzyme that specifically cleaves the RNA in an RNA:DNA hybrid. We employed a transgenic zebrafish system for inducible expression of *RNASEH1* that we recently used to examine R-loops in *sf3b1* mutant zebrafish. To target nuclear RNA:DNA hybrids and avoid potential toxicity to mitochondria, we utilized a truncated version of human *RNASEH1* (*M27RNASEH1*) that lacks its mitochondrial localization signal, and expressed it under a heat-shock-inducible promoter to allow for temporally-controlled expression [Tg(*hsp70:M27RNASEH1-GFP)*]. To decipher the impact of *ddx41* deficiency on R-loop levels in hematopoietic cells, we measured RNA:DNA hybrid immunofluorescence in purified *cd41:gfp*$^{+}$ hematopoietic stem and progenitor cells (HSPCs) from *ddx41* mutants and siblings. Consistent with the results in embryonic cells, we observed significantly elevated RNA:DNA hybrid levels in *ddx41* mutant HSPCs (**Fig. 1**). *These studies from year one strongly support a role for Ddx41 in preventing deleterious R-loop accumulation in vivo.*

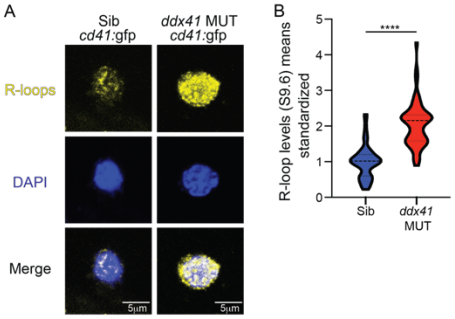


Fig 1. A. Confocal images showing immunofluorescence of R-loops and nuclei (DAPI) in sorted *cd41:*gfp^{+} HSPCs isolated from 40 hpf siblings (left) and *ddx41* mutants (right). Scale bar=5 μm. **B**. Quantification of R-loop levels from (A). Graph displays mean ± standard deviation (std) with p-value calculated with unpaired Student's t-test, ****=p<0.0001. N=3 replicates per experiment.

Aim 1.2: To determine if R-loops contribute to hematopoietic defects in *ddx41* mutants. Using the *M27RNASEH1-GFP* transgene, we explored the impact of excess R-loops on hematopoietic phenotypes in *ddx41* mutants. In our preliminary studies, we demonstrated that *ddx41* mutants have excessive HSPCs, but diminished erythrocytes, macrophages and neutrophils compared to control siblings. As quantified by assessment of both absolute number of *runx1:*mcherry^{+} HSPCs per embryo and *runx1 in situ* hybridization staining at 40 hpf, HSPC levels were significantly reduced in *ddx41* mutants after transient global RNASEH1-GFP expression (**Fig. 2A-E**). To elucdiate if the impact of excess R-loops on HSPC expansion was cell-intrinsic, we injected *ddx41* mutants with either *ubi*:gfp (control) or *fli1:M27RNASEH1-GFP* which would deplete R-loops only in *fli1*$^{+}$ endothelial cells including those that give rise to HSPCs. As quantified by assessment of *runx1:*mcherry^{+} HSPCs per embryo at 40 hpf, HSPC levels were significantly reduced in *ddx41* mutants injected with *fli1:M27RNASEH1-GFP* compared to control (**Fig. 2 F-G**). These data show that aberrant R-loop accumulation can intrinsically promote HSPC expansion *in vivo*.

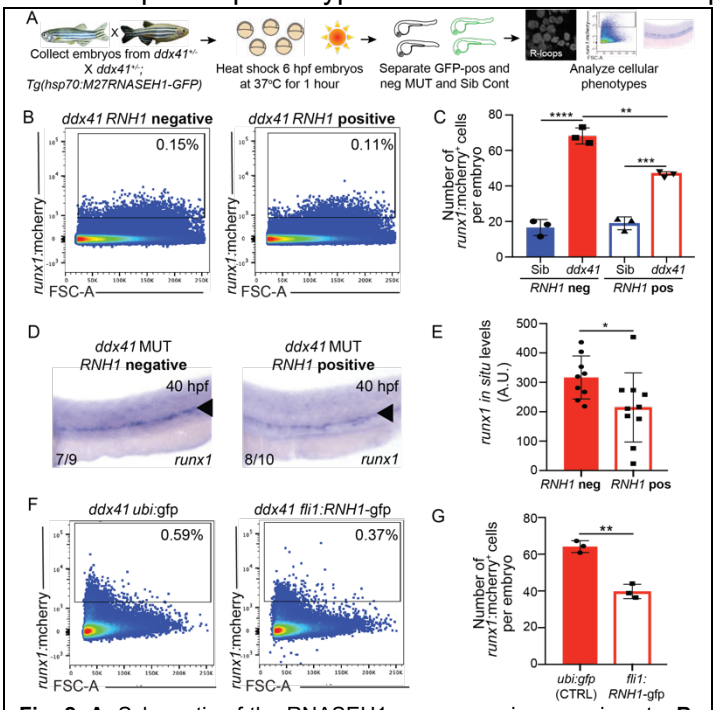


Fig. 2. A. Schematic of the RNASEH1 overexpression experiments. **B**. Flow cytometry plots of *runx1*:mcherry^{+} HSPCs from *Tg(hsp70:M27RNASEH1-GFP)*-negative (left) and *Tg(hsp70:M27RNASEH1-GFP)*-positive (right) *ddx41* mutants at 40 hpf. **C**. Graph depicting the number of *runx1*:mcherry^{+} HSPCs per embryo from (B). **D**. *In situ* hybridization of the HSPC marker *runx1* at 40 hpf in *Tg(hsp70:M27RNASEH1-GFP)*-negative *ddx41* mutants (left) and *Tg(hsp70:M27RNASEH1-GFP)*-positive *ddx41* mutants (right). Numbers on bottom left corner indicate the fraction of embryos with the same phenotype as the one depicted in the image. A higher magnification view of the AGM region is displayed. The aorta is marked with arrowheads. **E**. Quantification of *runx1 in situ* hybridization levels from (F). Quantification was done using Fiji. A.U.-arbitrary unit. **F.** Flow cytometry plots of *runx1*:mcherry^{+} HSPCs *ddx41* mutants injected with *ubi*:gfp (left, control) and *fli1:M27RNASEH1-GFP* (right) at 40 hpf. **G**. Graph depicting the number of *runx1*:mcherry^{+} HSPCs per embryo from (F). Graphs display means ± stds with p-values calculated with a one-way ANOVA with Tukey's multiple testing correction (C) or unpaired Student's t-test (E and G), *=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001, ns=not significant (p>0.05). N=3-6 replicates per experiment.

Paradoxically, when we transiently overexpressed *RNASEH1-GFP,* rather than suppressing the erythrocyte defect in *ddx41* mutants it exacerbated the anemia as measured by a decrease in the amount of hemoglobinized erythrocytes. *Combined, these data from year one demonstrated that proper R-loop balance is critical for hematopoietic homeostasis.*

Aim 1.3: To determine how R-loops regulate hematopoietic gene expression. R-loops are known to regulate gene expression via modulation of transcriptional elongation and epigenetic modifiers. Based on this, we hypothesized that DDX41-mediated R-loop regulation controls expression of critical genes for hematopoiesis, and that disruption of this regulation

contributes to hematopoietic dysfunction in *ddx41* mutants. As our initial results regarding R-loop regulation of hematopoiesis were in erythrocytes, we first explored R-loop regulation of erythroid gene expression. This hypothesis was supported by a recent study that mapped the genomic locations of R-loops in human K562 erythroleukemia cells and identified an R-loop near the promoter of the *GATA1* gene. To identify R-loop-regulated gene expression in *ddx41* mutant erythrocytes, we performed RNA-seq on cells isolated from *ddx41* mutants and siblings either with or without R-loop depletion via RNASEH1-GFP overexpression. From our initial analysis, we observed few changes in transcript expression levels, but significant alterations in splicing. There were more than 600 RNASEH1-sensitive alternative splicing events in *ddx41* mutant cells. A deeper characterization of this dataset is still underway.

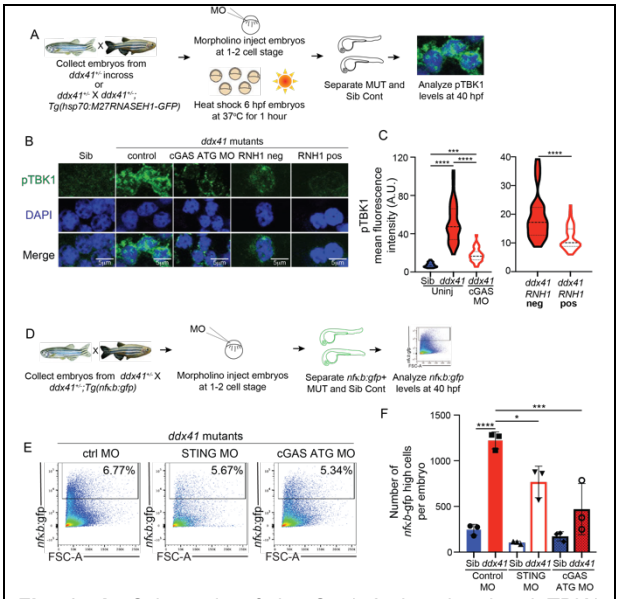


Fig. 3. A. Schematic of the Ser172-phosphorylated TBK1 (pTBK1) immunofluorescence experiments. **B**. Confocal images showing immunofluorescence of pTBK1 and nuclei (DAPI) in cells isolated from 40 hpf siblings, *ddx41* mutants (control), *ddx41* mutants injected with cGAS translation-blocking (ATG) MO, RNH1-negative *ddx41* mutants and RNH1-positive *ddx41* mutants. Scale bar=5μm. **C**. Quantification of pTBK1 levels from (B). **D**. Schematic of the cGAS-STING pathway knock-down experiments with *nf□b*:gfp quantification. **E**. Flow cytometry plots of *nfκb:*gfp$^+$ cells from *ddx41* mutants injected with control morpholino (left), splice-blocking *sting* morpholino (middle), or *cgas* ATG morpholino (right) at 40 hpf. **F**. Graph depicting the absolute number of *nfκb:*gfp-high cells per embryo shown in E . Graphs display means ± stds with p-values calculated with unpaired Student's t-test (C) or a one-way ANOVA with Tukey's multiple testing correction (F), *=p<0.05, ***=p<0.001, ****=p<0.0001. N=3-5 replicates per experiment.

Aim 2.1: To determine if inflammatory signaling is perturbed in *ddx41* mutants. Growing evidence suggests that perturbed inflammatory signaling is a common feature of MDS that contributes to disease pathophysiology. In addition to a role in RNA:DNA hybrid resolution, DDX41 is implicated in inflammatory signaling. DDX41 acts as an intracellular DNA sensor that activates the STING (Stimulator of Interferon Genes) pathway in both mammals and zebrafish. When activated, STING triggers a signaling cascade leading to activation of TBK1 (Tank Binding Kinase 1), the transcription factors NFκB (Nuclear Factor κ-B) and IRF3 (Interferon Response Factor 3), and production of Type I Interferon. Activation of the STING pathway in mice can induce HSPC cell cycle entry and mobilization, and thus could be involved in HSPC expansion. STING is activated by a number of cytosolic sensors in addition to DDX41. A recent study found that RNA:DNA hybrids can activate cGAS (cyclic GMP-AMP synthase), another cytosolic nucleic-acid sensor that triggers STING activation. To assess pathway activation in *ddx41* mutants and sibling controls, we measured phosphorylation of TBK1-Ser172 and NFκB transcriptional activity. We found elevated phospho-TBK1 and NFκB transcriptional reporter activity in *ddx41* mutants compared to siblings (**Fig. 3**). Pathway activation was diminished upon cGAS/STING knockdown and R-loop depletion. Studies examining interferon-responsive gene expression using the *interferon-phi:mcherry* transgenic zebrafish are still underway. *The data from year one of this project strongly indicate that excessive R-loops resulting from ddx41 loss trigger a cGAS/STING-mediated inflammatory response.*

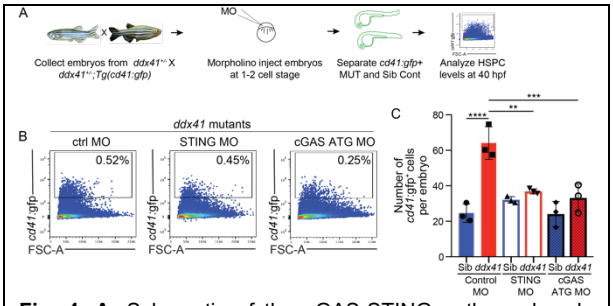


Fig. 4. A. Schematic of the cGAS-STING pathway knock-down experiments with HSPC quantification. **B**. Flow cytometry plots of *cd41:*gfp$^+$ HSPCs from *ddx41* mutants injected with control morpholino (left), splice-blocking *sting* morpholino (middle), or translation-blocking (ATG) *cgas* morpholino (right) at 40 hpf. **C**. Graph depicting the absolute number of *cd41:*gfp$^+$ cells per embryo shown in B. Graph displays means ± stds with p-values calculated with a one-way ANOVA with Tukey's multiple testing correction, **=p<0.01, ***=p<0.001, ****=p<0.0001, ns=not significant (p>0.05). N=3-6 replicates per experiment.

Aim 2.2: To determine if STING-mediated inflammatory signaling plays a role in hematopoietic defects in *ddx41* mutants. We tested the role of the cGAS/STING signaling components on the HSPC expansion in *ddx41* mutants (**Fig. 4A**). Lowering cGAS or STING activation using numerous genetic approaches significantly reduced HSPC numbers in *ddx41* mutants, as measured by quantification of *cd41:*gfp$^+$ HSPC numbers per embryo using flow cytometry and *runx1 in situ* hybridization (**Fig. 4B-C**). *These results indicate that the cGAS-STING inflammatory cascade mediates HSPC accumulation in ddx41 mutants*. Studies on the impact of inflammatory signaling in *ddx41* mutant erythrocyte, macrophage, and neutrophil defects is still underway.

Aim 2.3: To determine if elevated levels of R-loops alter inflammatory signaling in *ddx41* mutants. We deciphered that the elevated phospho-TBK1 levels were dampened following cGAS morpholino-mediated knockdown and following RNA:DNA hybrid depletion via RNASEH1-GFP overexpression (**Fig. 4**). Additionally, we demonstrated a decrease in some interferon-stimulated genes (ISGs) following R-loop depletion strengthening the model that R-loops activate a cGAS/STING inflammatory cascade leading to aberrant HSPC expansion

Professional development:
As part of my professional development, I attended the 2019 American Society of Hematology Conference in Orlando and 2020 International Society of Experimental Hematology Conference (Virtual).

Dissemination of results:
We have presented some of the data from this project at National and International Conferences (ASH and ISEH) as well invited seminar series at US-based research institutes.

Plans for next reporting period:
Experimental
Aim 1.1: To determine if DDX41 regulates R-loop levels- We have completed all the proposed experiments for this goal.
Aim 1.2: To determine if R-loops contribute to hematopoietic defects in *ddx41* mutants- In year two of the project, we will focus efforts on defining how R-loops impact erythrocyte, macrophage, and neutrophil number and differentiation.
Aim 1.3: To determine how R-loops regulate hematopoietic gene expression- In year one, we performed RNA-sequencing analysis of *gata1:gfp*$^{+}$ erythroid progenitors from *ddx41* mutants and siblings with or without RNASEH1-GFP overexpression to deplete R-loops. In year two, we will analyze the data to define how R-loops affect transcript levels and splice isoform selection in both mutants and siblings. We strive to identify signatures of R-loop regulated gene expression under normal (siblings) and defective (mutant) conditions. Additionally, as R-loops are known to impact chromatin, we will perform ATAC-sequencing in the same cell populations as the RNA-seq to decipher if R-loops alter chromatin accessibility in control or *ddx41* mutant erythroid progenitors. These data will provide insight into how homeostatic R-loops and excessive R-loops impact hematopoietic gene expression.
Aim 2.1: To determine if inflammatory signaling is perturbed in *ddx41* mutants- Our data from year one of the project demonstrated activation of cGAS/STING pathway in *ddx41* mutants. These data were primarily in embryonic cells. In year two, we will deduce if the pathway is activated intrinsically in hematopoietic cells by looking at phospho-TBK1 and NFkB reporter activity in HSPCs, erythrocytes, macrophages, and neutrophils.
Aim 2.2: To determine if STING-mediated inflammatory signaling plays a role in hematopoietic defects in *ddx41* mutants In year two of the project, we will focus on deciphering if cGAS/STING activity contributes to the diminished levels of *ddx41* mutant erythrocytes and myeloid cells.
Aim 2.3: To determine if elevated levels of R-loops alter inflammatory signaling in *ddx41* mutants- Our data from year one indicate that depletion of RNA:DNA hybrids in *ddx41* mutants lowers phospho-TBK1 levels. In year two, we will examine R-loop regulation of inflammatory gene expression. We will examine our RNA-seq data specifically for inflammatory genes as well as R-loop-regulated splicing alterations in *ddx41* mutant *gata1:gfp*$^{+}$ erythroid progenitors.

Publications
We currently have two manuscripts under considerations on this project. We anticipate these will be accepted during year two of the project.

4. Impact

Impact on the principal discipline: Mutations in *DDX41* in MDS were identified over 5 years ago, but how defects in this factor might contribute to hematopoietic dysfunction remains unknown. Our studies have revealed that *in vivo* DDX41 acts as a R-loop repressor and that loss of this function results in elevated R-loops, cGAS/STING pathway activation, and subsequent aberrant HSPC expansion. These mechanistic discoveries pave the way towards future biochemical studies on DDX41-RNA:DNA hybrid interactions, R-loop activation of cGAS, and the function of the cGAS/STING signaling in hematopoietic homeostasis. Moreover, these findings strongly suggest that inhibitors of the cGAS/STING pathway could be novel therapeutics for the treatment of MDS. Indeed, several STING antagonist are currently in pre-clinical and clinical develop for other diseases which could allow for more streamlined testing of these molecules for MDS.

Impact on other disciplines: Nothing to Report

Impact on society beyond science and technology: Nothing to Report

5. Changes/Problems
Nothing to Report

6. Products

Journal publications:
Joshua T. Weinreb, Noura Ghazale, Kith Pradhan, Varun Gupta, Kathryn S. Potts, Brad Tricomi, Noah J. Daniels, Richard A. Padgett, Sofia De Oliveira, Amit K. Verma, and Teresa V. Bowman[#]. Excessive R-loops Trigger an Inflammatory Cascade Leading to Aberrant HSPC Expansion. *Under review at Immunity.*

Joshua T. Weinreb, Elianna S. Pechman, Rachel Weil, Varun Gupta, and Teresa V. Bowman[#]. Ddx41 inhibition of DNA damage signaling permits erythroid progenitor expansion in zebrafish. *In revision at Haematologica.*

Conference papers and presentations:

Nov 2019	Five Points Seminar, New York Genome Center, New York, NY
June 2020	Blood and Bone Seminar, Virtual Seminar Series
Dec 2019	2019 American Society of Hematology Annual Meeting, Abstract Achievement Award and Oral Presentation, Orlando, FL
Aug 2020	2020 International Society of Experimental Hematology Meeting, Oral Presentation, Virtual Conference

7. Participants

Name:	Teresa V. Bowman, PhD
Project Role:	Principal Investigator
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	3
Contribution to project:	Dr. Bowman oversaw all of the work performed for the project, mentored and managed all participants, helped with data analysis, and wrote/edited all manuscripts.
Funding Support:	
Name:	Kathryn Potts, PhD
Project Role:	Postdoctoral Researcher
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	12
Contribution to project:	Dr. Potts performed the experiments in the project along with Joshua Weinreb.
Funding Support:	
Name:	Joshua Weinreb
Project Role:	Graduate Student
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	12
Contribution to project:	Joshua Weinreb performed all the experiments in the project along with Dr. Potts.
Funding Support:	Mr. Weinreb is supported by an F30 NIH predoctoral fellowship (1F30HL142161)
Name:	Varun Gupta
Project Role:	Bioinformatician
Research Identifier (e.g. ORCID ID):	

Nearest person month worked:	0.6
Contribution to project:	Varun Gupta performed all the computational analysis of the RNA-sequencing data.
Funding Support:	

Changes in active other support to Principal Investigator:
New Active Grants:
NIH/NIDDK: 1R01DK121738-01 2020-2024
"Crosstalk of Splicing and Signaling during HSPC Fate Choices"
Principal Investigator: Teresa Bowman

DOD/DTRA: 2020-2024
"Long non-coding RNAs in Venezuelan Equine Encephalitis Virus infection"
Principal Investigator: Steven Bradfute, University of New Mexico
Collaborator/Consultant: Teresa Bowman (salary only)

Completed Grants:
EvansMDS Foundation 2017-2019
"Finding New Therapies For MDS Through Chemical Screening In Zebrafish"
Principal Investigator: Teresa Bowman

NIH/NIDDK: 1R56DK121738-01 2019-2020
"Crosstalk of Splicing and Signaling during HSC Formation"
Principal Investigator: Teresa Bowman

Other organizations: Nothing to Report

8. Special Reporting Requirements: Nothing to Report

9. Appendices:

Conference Abstracts

ASH 2019 Conference Abstract:
Excessive R-loops Trigger an Inflammatory Cascade Leading to Aberrant Hematopoietic Stem and Progenitor Cell Expansion

Weinreb JT, Gupta V, and Bowman TV.

Mutations in *DEAD-box Helicase 41* (*DDX41*) are observed in patients with myelodysplastic syndromes (MDS) with inferior overall survival, implicating this factor in disease pathogenesis. DDX41 is an understudied factor with links to inflammation and RNA processing, but whose function in hematopoiesis is unknown. Using a novel zebrafish model of Ddx41-deficiency, we unveiled a critical role for Ddx41 in regulating hematopoietic stem and progenitor cells (HSPCs). Mutants for *ddx41* have more HSPCs than their siblings as measured by *in situ* hybridization for the HSPC marker *runx1* as well as quantification of *cd41:gfp*$^{+}$ and *runx1:mcherry*$^{+}$ HSPCs by flow cytometry. To uncover potential mechanisms driving HSPC expansion in *ddx41* mutants, we performed RNA-seq analysis of purified HSPCs from *ddx41* mutants and sibling controls. Expression of inflammatory target genes was significantly elevated in *ddx41* mutants including those regulated by the transcription factor NFκB (nuclear factor kappa-light-chain-enhancer of activated B cells). As inflammatory signaling and NFκB signaling in particular is elevated in MDS, we next explored Ddx41 regulation of NFκB signaling. Using an *in vivo nfkb:gfp* reporter, we demonstrated that NFκB signaling was indeed elevated in *ddx41* mutants. DDX41 is implicated in the STING (STimulator of INterferon Genes) and TBK1 (TANK binding kinase 1) pathway that mediates type I interferon production in response to foreign or damaged endogenous DNAs. By lowering pathway activity via knockdown of *sting* or pharmacological inhibition of TBK1, we determined that the STING/TBK1 pathway triggered excessive NFκB activation in *ddx41* mutants. Moreover, the elevated inflammatory signaling was detrimental to normal HSPC homeostasis as lowering STING/TBK1 pathway activation significantly reduced HSPC expansion in *ddx41* mutants, as measured by *runx1 in situ* hybridization and *cd41:gfp* flow cytometric quantification. R-loops, nucleic acid structures consisting of RNA:DNA hybrids

and displaced ssDNAs, can activate the STING/TBK1 pathway via the nucleic acid sensor cGAS (cyclic GMP-AMP synthase). Recent proteomic analysis showed that DDX41 can bind R-loops. Via immunofluorescence quantification of RNA:DNA hybrid levels in *ddx41* mutants *versus* sibling controls, we revealed that Ddx41 acts as a suppressor of R-loop accumulation. Using a *RNASEH1-GFP* transgenic zebrafish line that permits inducible depletion of R-loops, we revealed that excess R-loops in *ddx41* mutants promote HSPC expansion as measured by *runx1 in situ* hybridization and *runx1:mcherry* flow cytometric quantification. Moreover, using RT-qPCR analysis of NFκB target genes, we determined that R-loop depletion dampened NFκB activation in *ddx41* mutants, suggesting the R-loop-mediated effects on HSPC numbers was via inflammatory signaling. We deciphered cGAS activity was critical for this effect as both *nfkb:gfp* reporter activity and HSPC levels in *ddx41* mutants were diminished following treatment with a pharmacological inhibitor of cGAS. Our data demonstrate that DDX41 insufficiency triggers an R-loop-mediated inflammatory cascade leading to aberrant HSPC expansion and is the first study to delineate a functional consequence for R-loops in HSPC biology. Furthermore, recent studies showed R-loop accumulation in other common forms of MDS suggesting our novel *in vivo* findings are more broadly applicable to the MDS pathogenesis.

ISEH 2020 Conference Abstract:
Excessive R-loops Trigger an Inflammatory Cascade Leading to Aberrant Hematopoietic Stem and Progenitor Cell Expansion

Weinreb JT, Ghazale N, Pradhan K, Gupta V, Potts KS, Tricomi B, Daniels NJ, Padgett RA, Oliveira SD, Verma AK, and Bowman TV.

Aberrant expansion of hematopoietic stem and progenitor cells (HSPCs) is a major hallmark of myelodysplastic syndromes (MDS), although the underlying molecular regulation remains poorly understood. *DEAD-box Helicase 41* (*DDX41*) is an understudied MDS-mutated factor whose function in hematopoiesis is unknown. Using an *in vivo* zebrafish model of Ddx41-deficiency, we unveiled a critical role for this helicase in regulating HSPCs. Mutants for *ddx41* have more HSPCs than their siblings as measured by *in situ* hybridization and quantification by flow cytometry. We determined that *in vivo* Ddx41 suppresses the accumulation of R-loops, nucleic acid structures consisting of RNA:DNA hybrids and ssDNAs whose equilibrium is essential for cellular homeostasis. Using both global and hematovascular-restricted expression of *RNASEH1-GFP* to deplete R-loops, we revealed that excess R-loops in *ddx41* mutants promote HSPC expansion. To uncover potential mechanisms driving HSPC expansion in *ddx41* mutants, we identified elevated inflammatory target gene expression in HSPCs from *ddx41* mutants compared to sibling controls using RNA-seq analysis. We established that excess R-loop levels induced by *ddx41* deficiency triggered TBK1 phosphorylation and NFκB activation via the cGAS-STING pathway. Moreover, R-loop-triggered HSPC expansion was cGAS/STING-dependent. In human cells with diminished *DDX41,* blast counts, R-loop levels, and inflammatory signaling were elevated, suggesting a conservation of mechanism. These findings delineate a functional consequence for R-loop imbalance in HSPC biology and establish an *in vivo* link between R-loops, inflammation, and hematopoiesis.

Excessive R-loops Trigger an Inflammatory Cascade Leading to Aberrant HSPC Expansion

Joshua T. Weinreb[1,2], Noura Ghazale[1,2], Kith Pradhan[1,3], Varun Gupta[4], Kathryn S. Potts[1,2], Brad Tricomi[2,4], Noah J. Daniels[5], Richard A. Padgett[5], Sofia De Oliveira[1,6], Amit Verma[1,2,7], and Teresa V. Bowman[#1,2,7]

1. Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, NY, USA
2. Gottesman Institute for Stem Cell Biology and Regenerative Medicine, Albert Einstein College of Medicine, Bronx, NY, USA
3. Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA
4. Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY, USA
5. Cardiovascular and Metabolic Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA
6. Department of Medicine (Hepatology) and Marion Bessin Liver Research Center, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA
7. Department of Medicine (Oncology), Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA

#To whom correspondence should be addressed:
Address: 1300 Morris Park Avenue, Bronx, NY 10461
Email: teresa.bowman@einsteinmed.org
Phone: (718) 430-4001
Fax: (718) 430-8567

Character count: 39,833

Summary word count: 151

Number of figures: 6 main + 6 supplemental + 2 supplemental video

Number of tables: 6 supplemental

Number of references: 75

SUMMARY

Aberrant expansion of hematopoietic stem and progenitor cells (HSPCs) is a major hallmark of myelodysplastic syndromes (MDS), although the underlying molecular regulation remains poorly understood. *DEAD-box Helicase 41* (*DDX41*) is an understudied MDS-mutated factor whose function in hematopoiesis is unknown. Using an *in vivo* model of Ddx41-deficiency, we unveiled a critical role for this helicase in regulating HSPCs. We determined that *in vivo* Ddx41 suppresses the accumulation of R-loops, nucleic acid structures consisting of RNA:DNA hybrids and ssDNAs whose equilibrium is essential for cellular homeostasis. Excess R-loop levels in *ddx41* mutants triggered an inflammatory cascade via the activation of the cGAS-STING pathway that promoted aberrant HSPC expansion. Elevated R-loop accumulation and inflammatory signaling were observed in human cells with decreased DDX41, suggesting possible conservation of mechanism in humans. These findings delineate a functional consequence for R-loop imbalance in HSPC biology and establish an *in vivo* link between R-loops, inflammation, and hematopoiesis.

Key words: myelodysplastic syndromes, hematopoietic stem and progenitor cell, Ddx41, zebrafish, R-loops, inflammatory signaling

INTRODUCTION

The hematopoietic system is maintained by multipotent hematopoietic stem and progenitor cells (HSPCs) that generate all mature blood cells. HSPC dysfunction underlies myelodysplastic syndromes (MDS), a clonal hematopoietic disorder diagnosed by ineffective hematopoiesis and peripheral blood cytopenias, with a proclivity for malignant transformation (Lindsley and Ebert, 2013). Paradoxically, in MDS patients, there is a clonal selection/expansion of defective HSPCs over healthy cells (Sperling et al., 2017). The factors that contribute to aberrant HSPC accumulation remain poorly understood.

Deregulated inflammation is strongly implicated in MDS pathogenesis (Barreyro et al., 2018; Sallman and List, 2019). Aging-associated alterations to the inflammatory milieu contribute to changes in HSPC proliferation and differentiation (Pietras, 2017). Autoimmune and inflammatory conditions are encountered in approximately 25% of MDS patients, and correlate with an overall adverse patient outcome (Montoro et al., 2018). Animal models with elevated inflammatory signaling often present with MDS-like hematopoietic abnormalities, such as anemia and HSPC dysfunction (Fang et al., 2018; Schneider et al., 2016; Starczynowski et al., 2010; Varney et al., 2015). These findings indicate that balanced inflammatory regulation is crucial for healthy hematopoiesis and HSPC homeostasis.

Excessive R-loops, nucleic acid structures consisting of an RNA:DNA hybrid and ssDNA, have been observed in cells carrying MDS-associated mutations in splicing factors, although the contribution of these structures to hematopoietic dysfunction is unclear (Chen et al., 2018; Nguyen et al., 2018). In human disease, R-loop dysregulation is linked to cancer and mutations in R-loop-regulating factors are found in autoimmune disorders (Gunther et al., 2015; Lim et al., 2015). This clinical link suggests R-loop imbalance could derail normal inflammatory signaling. In the context of MDS, this could mean that aberrant R-loop accumulation could be mechanistically connected to altered inflammatory signaling. Despite the observation of R-loop imbalance in MDS and autoimmune

disorders, evidence for an *in vivo* consequence of this on the hematopoietic system is lacking.

Here, we show DEAD-box Helicase 41 (DDX41) sits at a crossroad of R-loop metabolism and inflammation to regulate HSPC homeostasis. In humans, *DDX41* is located in a commonly deleted region on chromosome 5q and is also mutated in both germline and acquired forms of MDS, suggesting a critical role for this helicase in the hematopoietic system (Polprasert et al., 2015; Sperling et al., 2017). DDX41 is associated with inflammation, splicing, and more recently R-loops, but its role in hematopoiesis is unclear (Parvatiyar et al., 2012; Polprasert et al., 2015; Wang et al., 2018; Zhang et al., 2013; Zhang et al., 2011). To uncover the *in vivo* function of DDX41 in HSPC biology, we established a zebrafish *ddx41* loss-of-function mutant, which is one of the first vertebrate animal models to study the *in vivo* requirements of Ddx41. We demonstrated that Ddx41 regulates HSPC number and that loss of *ddx41* leads to R-loop accumulation triggering an inflammatory response. This excess inflammation causes abnormal HSPC expansion via the cyclic GMP-AMP Synthase (cGAS)-Stimulator of Interferon Genes (STING) signaling pathway. Elevated R-loop accumulation and inflammatory signaling were also observed in human cells with decreased DDX41, suggesting possible conservation of mechanism in humans. Together, our data establish that Ddx41 acts as a negative regulator of HSPC expansion by maintaining R-loop homeostasis and preventing aberrant inflammatory signaling. These findings delineate that *in vivo* R-loop imbalance causes hematologic defects and provide a new mechanism for how R-loops could contribute to MDS pathogenesis.

RESULTS

Ddx41 regulates HSPC number.

We explored the *in vivo* role of DDX41 in hematopoiesis in a zebrafish *ddx41* loss-of-function mutant (*ddx41*^{sa14887}) (Kettleborough et al., 2013). DDX41 is highly conserved between humans and zebrafish with an overall 89% identity and 94% similarity between the two proteins with almost identical DEAD-Box and helicase domains (Figure S1A),

63 suggesting lessons learned about the *in vivo* role of zebrafish Ddx41 function in
64 hematopoiesis will be relevant to human DDX41.
65
66 In zebrafish, definitive hematopoiesis begins ~24 hours post fertilization (hpf), creating
67 HSPCs that can self-renew and reconstitute all adult blood lineages (Nik et al., 2017).
68 MDS is an HSPC disease, thus we examined HSPC levels in *ddx41* mutants by
69 measuring the expression of the HSPC marker *runx1* (Burns et al., 2002) via *in situ*
70 hybridization. At 40 hpf, levels of *runx1*$^{+}$ HSPCs were elevated in *ddx41* mutants
71 compared to sibling controls (Figures 1A-B). We also demonstrated this increase by
72 measuring the absolute number of HSPCs per embryo using flow cytometry quantification
73 with the HSPC fluorescent reporters *runx1:*mcherry (Tamplin et al., 2015) and *cd41:*gfp
74 (Lin et al., 2005). In agreement with the *runx1 in situ* results at 40 hpf, we observed a
75 significant increase in the number of both *runx1:*mcherry^{+} and *cd41:*gfp^{+} HSPCs in *ddx41*
76 mutants (Figures 1C-F). Taken together, our results indicate that *ddx41* mutants have a
77 significant increase in HSPCs.
78
79 We next wanted to elucidate if HSPC expansion in *ddx41* mutants is occurring via an
80 increase in proliferation and/or production. To examine proliferation, we analyzed cell
81 cycle status of *ddx41* mutant *cd41:*gfp^{+} HSPCs by flow cytometry quantification of DNA
82 synthesis via 5-Ethynyl-2'-deoxyuridine (EdU) incorporation and DNA content via DAPI
83 incorporation. At 40 hpf, *ddx41* mutant *cd41:*gfp^{+} HSPCs showed a reduction of cells in
84 S phase and an accumulation in the G2/M phase compared to sibling controls (Figures
85 S1C-D). These results are more in-line with a decrease in proliferation rather than an
86 increase indicating that HSPC expansion in *ddx41* mutants is not via increased cell
87 division.
88
89 A second mechanism we tested was increased HSPC production. During development,
90 HSPCs derive directly from the aortic hemogenic endothelium, located in the ventral floor
91 of the dorsal aorta, in a region known as the aorta-gonad-mesonephros (AGM), via a
92 process termed endothelial-to-hematopoietic transition (EHT) (Bertrand et al., 2010;
93 Boisset et al., 2010; Kissa and Herbomel, 2010; Ottersbach, 2019). This process involves

the transition of a subset of endothelial cells into HSPCs through an intermediate hemogenic endothelial cell state expressing features of both endothelial cells and HSPCs. First, we examined the vasculature via *in situ* hybridization for *flt4* (vein) and *notch1b* (aorta) and observed no difference between siblings and *ddx41* mutants, suggesting artery-vein specification was unaffected (Figures S1E-F). Next, we performed flow cytometry to quantify endothelial cells (*kdrl:*dsred$^+$*)*, hemogenic endothelial cells (*cd41*:gfp$^+$;*kdrl:*dsred$^+$*)*, and HSPCs (*cd41*:gfp$^+$) in sibling controls versus *ddx41* mutants. At 40 hpf, we observed no significant difference in the number of *kdrl:*dsred$^+$ endothelial cells, but a significant increase in the number of *cd41*:gfp$^+$ HSPCs and *cd41*:gfp$^+$;*kdrl:*dsred$^+$ double positive hemogenic endothelial cells in *ddx41* mutants (Figures 1G-H, S1G-H). To visualize EHT changes, we performed whole mount time-lapse confocal imaging of *cd41*:gfp$^+$;*kdrl:*dsred$^+$ sibling controls and *ddx41* mutants between ~36-44 hpf (Supplementary Videos 1-2). Consistent with the flow cytometry data, we observed that *ddx41* mutants had an increased number of *cd41*:gfp$^+$;*kdrl:*dsred$^+$ cells undergoing EHT morphological changes earlier in development than sibling controls. Taken together, our results indicate that the expansion of HSPCs in *ddx41* mutants is via an increased rate of endothelial-to-hematopoietic transition.

The control of R-loop levels by Ddx41 is critical for HSPC homeostasis.

It was recently demonstrated that cells carrying MDS-associated mutations in splicing factors have elevated R-loop levels and that removal of these R-loops can partially correct growth defects in hematopoietic progenitors (Chen et al., 2018; Nguyen et al., 2018). However, these studies were performed *ex vivo* outside of the normal, complex *in vivo* hematopoietic milieu. Cells possess numerous safeguards to prevent R-loop formation, such as splicing factors, as well as factors that remove them including ribonucleases that specifically cleave the RNA moiety of the RNA:DNA hybrid and helicases that can resolve RNA:DNA hybrids (Crossley et al., 2019). Based on the evidence that DDX41 has helicase activity (Yoneyama-Hirozane et al., 2017) and binds to RNA:DNA hybrids (Stavrou et al., 2018; Wang et al., 2018) along with the recent connection between R-loops and MDS, we hypothesized that Ddx41 might prevent R-loop accumulation *in vivo*.

To address this question, we measured R-loop levels in *ddx41* mutants and siblings. We performed immunofluorescence with the S9.6 antibody, which is a monoclonal antibody that specifically binds RNA:DNA hybrids, and is a mainstay in the field for R-loop identification (Boguslawski et al., 1986). To confirm the S9.6 signal in zebrafish *ddx41* mutants was specific to RNA:DNA hybrids, we used Tg(*hsp:M27RNASEH1-GFP*) zebrafish that express human *RNASEH1* fused to GFP under a heat shock-inducible promoter (Sorrells et al., 2018). RNase H1 is an enzyme which removes R-loops through the selective degradation of RNA in RNA:DNA hybrids and is the most rigorous confirmation to demonstrate R-loop dependency of a given phenotype (Paulsen et al., 2009; Suzuki et al., 2010). We induced RNASEH1-GFP expression by heat shocking *ddx41* mutants and siblings at 6 hpf, separated transgenic and non-transgenic siblings according to GFP expression, and then assessed R-loops levels (Figure 2A). We showed that R-loop levels were increased nearly two-fold in *ddx41* mutants compared to sibling cells and that the effect was suppressed by RNASEH1-GFP expression (Figures 2B-C). These data demonstrate that Ddx41 regulates R-loop levels *in vivo*.

To ascertain if elevated R-loop levels contributed to the HSPC expansion in *ddx41* mutants, we measured HSPC numbers following R-loop removal. As quantified by assessment of both absolute number of *runx1:*mcherry^{+} HSPCs per embryo and *runx1 in situ* hybridization staining at 40 hpf, HSPC levels were significantly reduced in *ddx41* mutants after RNASEH1-GFP expression (Figures 2D-G). This shows that excess R-loop levels lead to an *in vivo* HSPC expansion.

Finally, we assessed if excessive R-loops contributed to aberrant HSPC expansion via a cell-intrinsic or cell-extrinsic mechanism. To address this, we isolated *cd41:*gfp^{+} HSPCs from 40 hpf sibling controls and *ddx41* mutants using FACS and performed immunofluorescence with the S9.6 antibody to assess R-loop levels. Similar to the results in the whole embryo, R-loop levels were increased nearly two-fold in *ddx41* mutants *cd41:*gfp^{+} HSPCs compared to sibling *cd41:*gfp^{+} HSPCs (Figure S2). To elucdiate if the impact of excess R-loops on HSPC expansion was cell-intrinsic, we injected *ddx41* mutants with either *ubi*:gfp (control) or *fli1:M27RNASEH1-GFP* which would deplete R-

loops only in *fli1*$^+$ endothelial cells including those that give rise to HSPCs (Lawson and Weinstein, 2002). As quantified by assessment of *runx1*:mcherry$^+$ HSPCs per embryo at 40 hpf, HSPC levels were significantly reduced in *ddx41* mutants injected with *fli1:M27RNASEH1-GFP* compared to control (Figures 2H-I). These data show that aberrant R-loop accumulation can intrinsically promote HSPC expansion *in vivo*.

R-loops promote inflammatory gene expression in *ddx41* mutants.

To understand the mechanism behind the HSPC defect in *ddx41* mutants, we conducted RNA-sequencing (RNA-seq) on *cd41*:gfp$^+$ HSPCs isolated from *ddx41* mutants and siblings at 40 hpf. Of the nearly 19,000 genes analyzed, 1173 were downregulated and 1327 were upregulated in *ddx41* mutants compared to siblings (Figure 3A, Table S1).

To gain insight into the affected pathways, we performed gene set analysis on the upregulated and downregulated gene lists by comparing each to the Molecular Signature Database (MSigDb), a platform that computes overlaps between experimentally-derived gene lists and lists of genes in known pathways (Liberzon et al., 2015; Subramanian et al., 2005). In the downregulated list, genes associated with MTORC1 signaling, epithelial-mesenchymal transition, coagulation, and apical junctions were enriched (Figure 3B, Table S2). In the upregulated gene list, Spliceosome was the top gene set with Heme metabolism, DNA repair, and Interferon-alpha (IFN-α) response also enriched (Figure 3C, Table S3).

Ddx41 interacts with components of the spliceosome and is implicated in pre-mRNA splicing (Polprasert et al., 2015), thus we also examined how *ddx41* loss affected mRNA splicing in HSPCs. A total of 2,009 alternative splicing events were observed between *ddx41* mutants and siblings, with exon skipping being the most frequently altered splicing event followed by intron retention (Figures S3A-B, Table S4). These splicing defects are similar to those found by in *DDX41*-deleted and *DDX41*-mutated human cells (Polprasert et al., 2015). We were also able to detect other splicing defects not reported previously including alternative 5'-splice site usage, alternative 3'-splice site usage, and changes in mutually exclusive exon usage in *ddx41* mutants as compared to siblings. To understand

187 if particular pathways were enriched in the alternatively spliced gene list, we compared it
188 to the MSigDB and determined that immune system pathways were among the top enrich
189 pathways (Figure S3C). Among the alternatively spliced factors were several NFκB
190 pathway components including *map3k7*, *nfkb2*, *nfkbiab*, and *irak3*. These results depict
191 that Ddx41 influences inflammatory gene expression in HSPCs at multiple levels.
192
193 Type I interferons and their target genes were significantly elevated in *ddx41* mutant
194 HSPCs (Figure 3C, Table S3). We posited that excess R-loops might contribute to
195 inflammatory signaling. The cytosolic nucleic acid sensor cGAS binds to and can be
196 activated by RNA:DNA hybrids (Mankan et al., 2014). cGAS activity can induce a STING
197 signaling cascade resulting in increased type I IFN-responsive gene expression (Ablasser
198 et al., 2013; Li et al., 2013). To determine if R-loops contributed to inflammatory gene
199 expression, we measured the expression of *ifn-phi1*, which is a zebrafish type I IFN, and
200 several type I IFN-responsive genes identified in our RNA-seq analysis using RT-qPCR
201 in embryos ± RNASEH1-GFP expression. Depletion of R-loops significantly reduced
202 expression of many of these genes in *ddx41* mutants, demonstrating that excessive R-
203 loops promote inflammatory signaling (Figure 3D).
204
205 **Ddx41 regulates inflammatory signaling via the cGAS-STING pathway.**
206 Our model is that excess R-loops promote inflammation by activating the cGAS-STING
207 pathway. We first assessed if cGAS and R-loops were in the same cellular compartment,
208 and thus capable of activating cGAS-STING signaling as previously characterized
209 (Mankan et al., 2014). Prior studies showed that in addition to being in the cytoplasm,
210 cGAS can also be found in the nucleus, suggesting it can bind to nuclear RNA:DNA
211 hybrids (Gentili et al., 2019; Jiang et al., 2019; Liu et al., 2018; Volkman et al., 2019).
212 Consistent with these prior reports, we found cGAS to be predominantly in the nucleus of
213 zebrafish cells by immunofluorescence (Figure S4A-E). Additionally, R-loops were
214 present in both cellular compartments of *ddx41* mutant embryos (Figures 2B-C, S4F-G).
215 In combination, these results demonstrate that interactions between R-loops and cGAS
216 are possible in multiple cellular compartments.
217

218 Next, we ascertained if loss of *ddx41* led to activation of the cGAS-STING pathway.
219 Phosphorylation of Ser172 in Tank Binding Kinase 1 (TBK1) is induced downstream of
220 cGAS-STING activation, thus we measured phospho-TBK1 levels in sibling controls and
221 *ddx41* mutant cells by immunofluorescence and western blots (Bai et al., 2017). Levels
222 of phospho-TBK1 were significantly elevated in *ddx41* mutants and dependent on cGAS
223 and elevated R-loop levels (Figures 4A-C, S4H-I). These data indicate that the cGAS-
224 STING pathway is elevated when *ddx41* levels are diminished.
225
226 As stimulation of the cGAS-STING pathway ultimately results in activation of the
227 transcription factors IRF3/7 (interferon response factors 3 and 7) and NFκB (nuclear
228 factor κ light-chain enhancer of activated B cells), we next measured pathway activity
229 using an NFκB activity fluorescent reporter zebrafish that expresses GFP under the
230 transcriptional control of a NFκB-responsive promoter (*nfκb:*gfp) (Kanther et al., 2011).
231 We quantified the absolute number of cells with high *nfκb:*gfp reporter activity using flow
232 cytometry and demonstrated a significant increase in *ddx41* mutants (Figures 4D-E). We
233 also determined that NFκB activity was increased in *runx1:*mcherry^{+} HSPCs and
234 *kdrl:*dsred^{+} endothelial cells in *ddx41* mutants (Figures S4J-M). These data support that
235 Ddx41 regulates NFκB signaling in numerous cell types including HSPC and endothelial
236 cells.
237
238 Next, we determined if inhibiting components of the cGAS-STING pathway would dampen
239 inflammatory signaling in *ddx41* mutants by measuring *nfκb:*gfp reporter activity (Figure
240 4F). Morpholino knockdown of both STING and cGAS lowered *nfκb:*gfp reporter activity
241 in *ddx41* mutants (Figures 4G-H, S4E,N-P). Taken together, these data indicate that
242 cGAS-STING pathway activation is functionally important for the inflammatory response
243 in *ddx41* mutants.
244
245 **Ddx41 regulates HSPC number via the cGAS-STING inflammatory pathway.**
246 Inflammatory signaling plays a physiological role in regulating HSPC biology from the
247 earliest stages of stem cell formation in the embryo and throughout the aging process
248 (Espin-Palazon et al., 2018; Pietras, 2017). In particular, NFκB signaling is required for

the specification of HSPCs in embryonic zebrafish hematopoiesis (Espin-Palazon et al., 2014; He et al., 2015). Based on our finding that Ddx41 regulates NFκB levels via the cGAS-STING axis, we tested the role of these signaling components on the HSPC expansion in *ddx41* mutants (Figure 5A). Lowering cGAS and STING levels using numerous genetic approaches significantly reduced HSPC numbers in *ddx41* mutants, as measured by quantification of *cd41:*gfp^{+} HSPC numbers per embryo using flow cytometry and *runx1 in situ* hybridization (Figures 5B-C, S5). These results indicate that the cGAS-STING inflammatory cascade mediates HSPC accumulation in *ddx41* mutants.

Next, we determined if the R-loop effect on HSPC accumulation was dependent or independent of the cGAS-STING signaling pathway. To do so, we assessed the impact of combined RNASEH1-GFP overexpression and STING knockdown on HSPC numbers in *ddx41* mutants (Figure 5D). If excess R-loops and STING are working via independent or parallel pathways, then we would expect an additive effect with further HSPC reduction compared to each manipulation alone. If a single pathway, then we would not see an additive suppression on HSPC expansion. We observed no difference between HSPC number in 40hpf *ddx41* mutants with STING knockdown alone or in combination with RNASEH1-GFP overexpression (Figures 5E-F). These data support a single pathway with R-loops impacting HSPCs via a STING-mediated pathway.

Low *DDX41* expression levels predict high blast counts and poor MDS prognosis.

In humans, *DDX41* is located in a commonly deleted region on chromosome 5q and is also mutated in both germline and acquired forms of MDS (Polprasert et al., 2015). Germline *DDX41* frameshift mutations lead to loss-of-function and somatic missense mutations in *DDX41* are thought to result in hypomorphic helicase activity (Polprasert et al., 2015; Yoneyama-Hirozane et al., 2017). Overall, *DDX41* mutations are rare in MDS, but since transcriptional alterations are common in MDS, we wanted to determine whether lower *DDX41* expression was associated with poor clinical outcomes in this disease. Since MDS are characterized by peripheral blood cytopenias and increased bone marrow myeloblasts, we compared blood counts and bone marrow blast levels in a cohort of 183 MDS patients (Pellagatti et al., 2010). We stratified the patients based on *DDX41*

expression in $CD34^+$ bone marrow cells comparing hematopoietic parameters in low (bottom 20^{th} percentile) *versus* high (top 20^{th} percentile)-expressing patients (Figure S6A). Reduced WBC, neutrophil, and platelet counts were seen in MDS cases with lower levels of *DDX41* (Figures S6B-D). Significantly, higher bone marrow blast counts were found in patients with lower *DDX41* levels (Figure 6A). Moreover, patients with low *DDX41* expression had inferior overall survival compared to patients with high levels of *DDX41* (Figure 6B). These data are suggestive that in addition to the rare *DDX41* mutations, lower *DDX41* expression could contribute to MDS pathogenesis.

Using this same patient dataset, we determined if lower *DDX41* levels in human $CD34^+$ HSPCs was also concordant with elevated inflammatory gene expression. Since we observed differences in clinical features between MDS samples with low *versus* high *DDX41* expression, we asked if the genes associated with varying *DDX41* levels were also predictive of differences in hematopoietic parameters. We defined the genes with elevated expression in *DDX41* low *versus* high samples (Table S5), and then performed gene set comparison to REACTOME pathways. Similar to our findings in *ddx41* mutant zebrafish, we observed that immune system pathways were among the top enriched (Figure 6C). We then took the genes that were highly expressed in *DDX41* low MDS samples and within the 'Innate immune system' pathway and assessed how their expression correlated with blast counts in the cohort of 183 MDS patients (Figure 6D). We observed a strong correlation with expression of these genes with high bone marrow blast counts (Figure 6E). These data support that low expression of *DDX41* in human $CD34^+$ HSPCs leads to increased inflammatory signaling that correlates with increased blast levels.

Finally, we assessed if *DDX41* hypomorphic mutations could modulate R-loop levels. We focused on the $DDX41^{R525H}$ mutant which is a common hotspot somatic mutation that is thought to result in hypomorphic helicase activity (Polprasert et al., 2015; Yoneyama-Hirozane et al., 2017). We analyzed R-loop levels in HEK-293 cells overexpressing wildtype DDX41 (control) or mutant $DDX41^{R525H}$. We showed that R-loop levels were significantly increased in $DDX41^{R525H}$-overexpressing cells compared to $DDX41^{wildtype}$-

overexpressing cells (Figures 6F-G). These data demonstrate that similar to our finding in zebrafish, human DDX41 regulates R-loop levels.

DISCUSSION

Although *DDX41* mutations have been identified in MDS, its function in hematopoiesis or MDS pathology is unknown. Here, we unveil a critical role for Ddx41 as a key gatekeeper to maintain balanced inflammatory signaling, a necessary component for HSPC homeostasis. We demonstrated that deficiency of *ddx41* triggers inflammatory signaling via activation of a cGAS-STING pathway. Ultimately, we determined that R-loop dysregulation in *ddx41* mutants triggers this inflammatory cascade leading to aberrant HSPC expansion, illuminating a detrimental role for R-loop imbalance in HSPC biology.

Inflammatory signaling is well established to play a critical role in HSPC biology from birth to old age (Espin-Palazon et al., 2018; Pietras, 2017). During embryonic development, prior work suggested sterile inflammation, a non-infection-mediated activation of immunomodulatory signaling, is at the heart of the inflammatory signals required for HSPC formation (Espin-Palazon et al., 2014; He et al., 2015; Li et al., 2014; Sawamiphak et al., 2014). However, the upstream triggers of such inflammation remain unclear. Our work shows that endogenous signals, such as aberrant R-loops, can function as innate damage-associated molecular patterns (DAMPs) during development to alter HSPC formation. This finding opens new avenues for considering the role for other naturally occurring internal inflammatory triggers as native constituents of the HSPC formation microenvironment.

Previous work demonstrated that DDX41 can act as a sensor for cytoplasmic DNA and bacterial cyclic di-GMP/AMP activating a type I IFN immune response (Parvatiyar et al., 2012; Zhang et al., 2013; Zhang et al., 2011). Out of context, this data is in apparent contradiction to our findings that deficiency of *ddx41* results in increased type I IFN signaling. The previous studies were mainly performed in myeloid dendritic cells (mDCs) or monocytes, which may reflect how DDX41 functions in immune effector cells. DDX41 is ubiquitously expressed (Fagerberg et al., 2014). Zebrafish *ddx41* mutants have the

loss-of-function mutations in all cells, opening the possibility of a different role for DDX41 in non-immune cell types, such as HSPCs. As humans with *DDX41* germline mutations also carry the mutation in all cells, it is possible that this effect is critical in their development of MDS as well. Furthermore, our analysis shows that MDS CD34^{+} HSPCs with low levels of *DDX41* display increased inflammatory signaling, suggesting a conserved HSPC response to diminished *DDX41* levels. Another contributing factor to the differences between the prior immune signaling studies and our work is the pleiotropic functions of DDX41. Although, numerous cellular roles of DDX41 are described, how these processes interact and influence one another is still at its infancy. Our data indicate at least one example of how DDX41 regulation of R-loops and inflammatory signaling are intertwined.

Although aberrant R-loop accumulation has been suggested as a unifying mechanism to explain MDS pathogenesis in splicing factor mutant cells, *in vivo* evidence for a negative role for R-loop imbalance in hematopoiesis was lacking. Several groups showed that these excessive R-loops caused DNA damage and sensitized cells to DNA damaging agents (Chen et al., 2018; Nguyen et al., 2018; Sorrells et al., 2018). Our current work revealed that Ddx41 regulates R-loop levels, and that in addition to DNA damage, these aberrant structures also increase inflammatory signaling. The ultimate result of this signaling cascade is the anomalous expansion of HSPCs. Our findings do not preclude R-loop-mediated genomic instability as another contributor to pathologic hematopoiesis but do indicate that R-loop-mediated inflammation is a critical driver of aberrant HSPC expansion. Combined, these results imply that targeting both R-loop-mediated DNA damage and inflammatory responses could have a greater therapeutic value than either approach alone.

Elevated inflammatory signaling is a hallmark of MDS (Fang et al., 2018; Schneider et al., 2016; Starczynowski et al., 2010; Varney et al., 2015) with approximately 25% of cases linked to autoimmunity or other pro-inflammatory conditions (Montoro et al., 2018). R-loops are linked not only to cancer, but mutations in R-loop-regulating factors are also found in autoimmune disorders (Gunther et al., 2015; Lim et al., 2015). Though this clinical

373 linkage suggests R-loop imbalance could derail normal inflammatory signaling, our work
374 provides *in vivo* and mechanistic evidence of this occurring. Overall our data indicate that
375 R-loop-mediated inflammation might be an unappreciated mechanism driving the clinical
376 connections between autoimmune conditions and hematopoietic abnormalities in MDS.
377
378 Activation of the STING pathway was recently shown to have anti-leukemia effects
379 through a leukemia-extrinsic mechanism via the induction of dendritic and T-cell activation
380 resulting in tumor clearance (Curran et al., 2016). While this study seems contradictory
381 to our findings, there are several key differences to consider. These types of immune
382 effector cells do not arise until much later in zebrafish development; thus, dendritic or T-
383 cell responses could not be contributing to the HSPC phenotype in 40 hpf *ddx41* mutants.
384 Moreover, our data indicate that the effect of Ddx41-regulated R-loops on HSPC
385 expansion is cell intrinsic. These findings indicate that there are likely cell-type-specific
386 responses to *ddx41* loss and/or STING pathway activation. Moreover, other groups have
387 shown that activation of the cGAS-STING pathway can promote solid tumor initiation and
388 growth (Bakhoum et al., 2018; Lemos et al., 2016; Ng et al., 2018). All these data indicate
389 more complexity in STING-mediated effects then were previously appreciated.
390
391 Splicing mutations are commonly found in MDS (Haferlach et al., 2014; Yoshida et al.,
392 2011). DDX41 interacts with multiple components of the spliceosome (Polprasert et al.,
393 2015). Our work aligns with prior studies showing *DDX41* insufficiency associates with
394 numerous deleterious splicing outcomes. If and how these splicing events contribute to
395 MDS pathogenesis is uncertain. We show that immune signaling components are
396 commonly mis-spliced in *ddx41* mutants. Therefore, although we demonstrated a clear
397 impact of R-loops on inflammatory signaling in *ddx41* mutants, our data do not preclude
398 the possibility that splicing alterations could also contribute to perturbed inflammatory
399 signaling and HSPC expansion. As mentioned, MDS-associated mutations in splicing
400 factors, similar to our results with *ddx41* deficits, cause R-loop imbalance in addition to
401 splicing alterations. The crosstalk of pre-mRNA processing and R-loop regulation is a
402 major outstanding question in the field that remains to be addressed.
403

In sum, our study uncovers an important link between R-loops, inflammation, and the developing hematopoietic system. We demonstrate that loss of *ddx41* leads to increased R-loop levels, which trigger an inflammatory response contributing to HSPC expansion. This mechanistic connection between R-loops and inflammation gives us a greater understanding of HSPC development. Furthermore, combined with the observation of R-loop imbalance in other common forms of MDS, our findings provide a strong rationale for exploring R-loop-mediated inflammatory responses in other pre-clinical models of MDS and imply that our novel *in vivo* findings could be more broadly applicable to the MDS pathogenesis in general.

ACKNOWLEDGMENTS

This work was funded by American Cancer Society RSG-129527-DDC, Kimmel Foundation, the Evans MDS Foundation, DOD BM180109, and 1R56DK121738-01 (to TVB), MSTP training grant T32GM007288-45 and F30 fellowship 1F30HL142161 (to JTW), R01HL139487 and R01HL150832 (to AV), and the American Australian Association Sir Rupert Murdoch Postdoctoral Fellowship (to KSP).

We want to thank Eirini Trompouki, Charles Query, Andreas Jenny, and Sara Nik for helpful discussions on this work. We want to thank Mary Goll for sharing the *sting*^{mk30} mutant zebrafish. We also want to acknowledge the assistance of numerous core facilities at Albert Einstein College of Medicine including Flow Cytometry, Analytical Imaging, and Genomics Facilities (funded by NCI Cancer Grant P30CA013330), and the Zebrafish Core Facility.

AUTHORSHIP CONTRIBUTIONS

JTW, AV, and TVB designed the project experimental approach. JTW, NG, KSP, BT, and SDO performed the experiments. KP and VG performed bioinformatics analysis. NJD and RAP shared critical reagents. JTW and TVB analyzed the data. JTW and TVB wrote the manuscript; all authors reviewed and approved the manuscript.

DISCLOSURE OF COMPETING INTERESTS

The authors declare no competing financial interests.

436 **FIGURE LEGENDS**

437 **Figure 1. Ddx41 regulates HSPC number. A.** *In situ* hybridization of the HSPC marker
438 *runx1* at 40 hpf in sibling controls (left) and *ddx41* mutants (right). Numbers on bottom left
439 corner indicate the fraction of embryos with the same phenotype as the one depicted in
440 the image. Inset underneath shows a higher magnification (8X) view of the boxed AGM
441 region above. The aorta is marked with arrowheads. **B**. Quantification of *runx1 in situ*
442 hybridization levels from (A). Quantification was done using Fiji. A.U.-arbitrary unit. **C,E**.
443 Flow cytometry plots of *runx1*:mcherry$^+$ (C) and *cd41:*gfp$^+$ (E) HSPCs from sibling controls
444 (left) and *ddx41* mutants (right) at 40hpf. **D,F**. Graphs depicting the absolute number of
445 *runx1*:mcherry$^+$ (C) and *cd41:*gfp$^+$ (E) per embryo at 40 hpf. **G**. Flow cytometry plots of
446 *kdrl*:dsred$^+$, *cd41:*gfp$^+$, and *kdrl*:dsred$^+$; *cd41:*gfp$^+$ double positive cells from sibling
447 controls (left) and *ddx41* mutants (right) at 40hpf. **H**. Graphs depicting the absolute
448 number of *kdrl*:dsred$^+$; *cd41:*gfp$^+$ double positive hemogenic endothelial cells per embryo
449 at 40 hpf. Graphs display means ± standard deviations (stds) with p-values calculated
450 with unpaired Student's t-test, *=p<0.05, ***=p<0.001, ****=p<0.0001. N=3-5 replicates
451 per experiment.
452
453 **Figure 2. The control of R-loop levels by Ddx41 is critical for HSPC homeostasis.**
454 **A**. Schematic of the RNASEH1 overexpression experiments. **B**. Confocal images
455 showing immunofluorescence of R-loops and nuclei (DAPI) in cells isolated from 28 hpf
456 siblings (left) and *ddx41* mutants (right) that are either Tg*(hsp70:M27RNASEH1-GFP)*-
457 negative (left) or *Tg(hsp70:M27RNASEH1-GFP)*-positive (right). Scale bar=5μm. **C**.
458 Quantification of R-loop levels from (B). **D**. Flow cytometry plots of *runx1*:mcherry$^+$
459 HSPCs from *Tg(hsp70:M27RNASEH1-GFP)*-negative (left) and
460 *Tg(hsp70:M27RNASEH1-GFP)*-positive (right) *ddx41* mutants at 40 hpf. **E**. Graph
461 depicting the number of *runx1*:mcherry$^+$ HSPCs per embryo from (D). **F**. *In situ*
462 hybridization of the HSPC marker *runx1* at 40 hpf in *Tg(hsp70:M27RNASEH1-GFP)*-
463 negative *ddx41* mutants (left) and *Tg(hsp70:M27RNASEH1-GFP)*-positive *ddx41*
464 mutants (right). Numbers on bottom left corner indicate the fraction of embryos with the
465 same phenotype as the one depicted in the image. A higher magnification view of the
466 AGM region is displayed. The aorta is marked with arrowheads. **G**. Quantification of *runx1*

467 *in situ* hybridization levels from (F). Quantification was done using Fiji. A.U.-arbitrary unit.
468 **H.** Flow cytometry plots of *runx1*:mcherry^{+} HSPCs *ddx41* mutants injected with *ubi*:gfp
469 (left, control) and *fli1:M27RNASEH1-GFP* (right) at 40 hpf. **I**. Graph depicting the number
470 of *runx1*:mcherry^{+} HSPCs per embryo from (H). Graphs display means ± stds with p-
471 values calculated with a one-way ANOVA with Tukey's multiple testing correction (C and
472 E) or unpaired Student's t-test (G and I), *=p<0.05, **=p<0.01, ***=p<0.001,
473 ****=p<0.0001, ns=not significant (p>0.05). N=3-6 replicates per experiment.
474
475 **Figure 3. R-loops promote inflammatory gene expression in *ddx41* mutants.**
476 **A**. Volcano plot displaying differentially expressed genes between *cd41*:gfp^{+} HSPCs from
477 *ddx41* mutants and siblings. Significant differences are defined as FDR <0.05 and log2
478 fold change >1. Black vertical lines denote the fold-change threshold and the black
479 horizontal line denotes the FDR threshold. Three biological replicates for both *ddx41*
480 mutants and siblings were used to generate RNA-sequencing data. **B-C**. Representative
481 charts of pathways significantly enriched in genes downregulated (B) or upregulated (C)
482 in *ddx41* mutant HSPCs compared to sibling controls as determined by MSigDB analysis.
483 **D**. Graph of RT-qPCR analysis of the expression of Type I IFN responsive genes between
484 sibling controls and *ddx41* mutants that are *Tg(hsp70:M27RNASEH1-GFP)*-negative
485 versus *Tg(hsp70:M27RNASEH1-GFP)*-positive. Expression levels were normalized to β-
486 actin levels. All levels are relative to the RNASEH1-GFP (RNH1) negative sibling controls
487 expressed as ddCt values, with positive values reflecting higher expression and negative
488 values reflecting lower expression. Graph displays means ± standard error mean with p-
489 values calculated with an unpaired t-test, *=p<0.05. N=4 replicates per experiment.
490
491 **Figure 4. Ddx41 regulates inflammatory signaling via the cGAS-STING pathway.**
492 **A**. Schematic of the Ser172-phosphorylated TBK1 (pTBK1) immunofluorescence
493 experiments. **B**. Confocal images showing immunofluorescence of pTBK1 and nuclei
494 (DAPI) in cells isolated from 40 hpf siblings, *ddx41* mutants (control), *ddx41* mutants
495 injected with cGAS translation-blocking (ATG) MO, RNH1-negative *ddx41* mutants and
496 RNH1-positive *ddx41* mutants. Scale bar=5μm. **C**. Quantification of pTBK1 levels from
497 (B). **D**. Flow cytometry plots of *nfκb*:gfp^{+} cells from sibling controls (left) and *ddx41*

mutants (right) at 40 hpf. **E**. Graph depicting the absolute number of *nfκb:*gfp-high cells per embryo shown in (D). **F.** Schematic of the cGAS-STING pathway knock-down experiments with *nfκb:*gfp quantification. **G**. Flow cytometry plots of *nfκb:*gfp$^+$ cells from *ddx41* mutants injected with control morpholino (left), splice-blocking *sting* morpholino (middle), or *cgas* ATG morpholino (right) at 40 hpf. **H**. Graph depicting the absolute number of *nfκb:*gfp-high cells per embryo shown in G . Graphs display means ± stds with p-values calculated with unpaired Student's t-test (E) or a one-way ANOVA with Tukey's multiple testing correction (C and H), *=p<0.05, ***=p<0.001, ****=p<0.0001. N=3-5 replicates per experiment.

Figure 5. Ddx41 regulates HSPC number via the cGAS-STING inflammatory pathway. A. Schematic of the cGAS-STING pathway knock-down experiments with HSPC quantification. **B**. Flow cytometry plots of *cd41:*gfp$^+$ HSPCs from *ddx41* mutants injected with control morpholino (left), splice-blocking *sting* morpholino (middle), or translation-blocking (ATG) *cgas* morpholino (right) at 40 hpf. **C**. Graph depicting the absolute number of *cd41:*gfp$^+$ cells per embryo shown in B. **D**. Schematic of the STING knock-down and RNASEH1 overexpression experiment with HSPC quantification. **E.** Flow cytometry plots of *runx1*:mcherry$^+$ HSPCs from *Tg(hsp70:M27RNASEH1-GFP)*-negative (top left) and *Tg(hsp70:M27RNASEH1-GFP)*-positive (top right) uninjected *ddx41* mutants *versus* *Tg(hsp70:M27RNASEH1-GFP)*-negative (bottom left) and *Tg(hsp70:M27RNASEH1-GFP)*-positive (bottom right) STING morpholino-injected *ddx41* mutants at 40 hpf. **F**. Graph depicting the number of *runx1*:mcherry$^+$ HSPCs per embryo from (E). Graphs display means ± stds with p-values calculated with a one-way ANOVA with Tukey's multiple testing correction, **=p<0.01, ***=p<0.001, ****=p<0.0001, ns=not significant (p>0.05). N=3-6 replicates per experiment.

Figure 6. Conserved properties in *DDX41* low and mutated human cells.
A. Graph depicting bone marrow blast percentages of MDS patients with high (top 20%) or low (bottom 20%) *DDX41* expression. **B**. Kaplan-Meier survival curve comparing MDS patients with high or low *DDX41* expression. **C.** Representative charts of REACTOME pathways significantly enriched in *DDX41* low patients' CD34$^+$ HSPCs compared to

DDX41 high patients' $CD34^+$ HSPCs. **D.** Venn diagram showing overlap between genes that were highly expressed in *DDX41* low MDS samples and within the 'Innate immune system' pathway and assessed for how their expression correlated with blast counts in the cohort of 183 MDS patients in (E). **E.** Heatmap correlating clinical hematological variables with the genes in the REACTOME 'innate immune signature' pathway from *DDX41* low patients' $CD34^+$ HSPCs (D). Each cell represents the correlation between the 183 values of the clinical variable (column), and the 183 values of the gene probe (row) matched by patient. **F.** Confocal images showing immunofluorescence of R-loops and nuclei (DAPI) in HEK-293 cells that are overexpressing DDX41 wildtype (left) or a DDX41 R525H mutant (right). Scale bar=10μm. **G.** Quantification of R-loop levels from (F). Graphs display means ± stds with p-values calculated with unpaired Student's t-test, *=p<0.05, ****=p<0.0001.

STAR METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Teresa Bowman (teresa.bowman@einsteinmed.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Zebrafish

Zebrafish were maintained as described (Lawrence, 2011). All fish were maintained according to IACUC-approved protocols in accordance with Albert Einstein College of Medicine research guidelines. We utilized several established zebrafish lines. Genotyping and phenotyping to confirm their identity was performed for each new generation and for embryos. Mutants for *ddx41* (*ddx41*^{sa14887}) were acquired from the Zebrafish International Resource Center (ZIRC) as part of the Sanger Institute Zebrafish Mutation Project (Kettleborough et al., 2013). The mutation results in a premature stop codon at Tyrosine 410. As early as 24 hpf, there appears to be slight brain death and abnormal curvature of the tail in *ddx41* mutants, allowing for phenotypic identification of mutants compared to siblings based on morphology (Figure S1B). Homozygous mutants are embryonic lethal, dying between 2-3dpf. This allows for the assessment of definitive hematopoietic development in zebrafish up to 3 dpf but precludes the assessment of hematopoiesis beyond this point of development. Mutants for *sting* (*sting*^{mk30}) were acquired from Mary Goll (Rajshekar et al., 2018). The following strains were used: *cd41:*gfp (Lin et al., 2005), *runx1:*mcherry (Tamplin et al., 2015), *kdrl:*dsred (Nakayama et al., 2012), *nfκb:*gfp (Kanther et al., 2011), and *Tg(hsp70:M27RNASEH1-GFP)* (Sorrells et al., 2018). *ddx41* genotyping was determined by mutant-specific PCR, which was initially verified by Sanger sequencing. Additionally, the genotype of *ddx41*$^{sa14887/+}$ carriers were validated by paired mating of heterozygous adult animals followed by examination of embryos for the expected homozygous recessive phenotype. All transgenic strains used in this study were scored for their distinctive fluorescence pattern during embryogenesis. Genomic DNA for genotyping was isolated by alkaline lysis (Meeker et al., 2007). The *ddx41*^{sa14887} mutant

571 and wildtype-specific PCR products were analyzed by gel electrophoresis. Genotyping
572 primers are in the Key Resources Table S6.
573
574 **Cell Lines**
575 The parental human embryonic kidney 293 cells (HEK-293) cell line was acquired from
576 ATCC (CRL-1573). HEK-293 cells were grown on 100 mm plates in Dulbecco's Modified
577 Eagle's Medium with 4.5 g/L glucose, l-glutamine, and sodium pyruvate (DMEM, Corning
578 Cellgro #10-013-CV) supplemented with 10% (v/v) fetal bovine serum (Atlanta
579 Biologicals) and 100 U/mL penicillin–streptomycin (Gibco) at 37 °C and 5% (v/v) CO_2.
580 The DDX41-V5 lentiviral expression vector (pLX304, clone ID: HsCD00442077) was
581 obtained from DNASU Plasmid Repository. The DDX41^{R525H}-V5 plasmid was generated
582 by introducing a point mutation in the DDX41-V5 vector using QuikChange site directed
583 mutagenesis kit with the following primers (DDX41-R525H-F:
584 CGCACCGGGCACTCGGGAAAC and DDX41-R525H-R:
585 GCCAATCCGGTGTACATAGTTCTC). DDX41-V5 and DDX41^{R525H}-V5 vectors were
586 used to generate lentiviral supernatants in HEK-293T cells. HEK-293 cells were subjected
587 to 8ug/ul polybrene and viral supernatants for 24hrs followed by antibiotic selection using
588 10ug/ml of blasticidin to generate stable HEK-293 wildtype DDX41-V5 or mutant
589 DDX41^{R525H}-V5 overexpression cell lines. DDX41 protein levels were validated using V5
590 (A190-120A, Bethyl) and DDX41 (NB100-57488, Novus Biosciences) antibodies. These
591 constructs are published in (Polprasert et al., 2015).
592
593 **METHOD DETAILS**
594 **Zebrafish embryo morpholino and construct injections**
595 Embryos derived from *ddx41* heterozygous incrosses were injected with 1.5 ng
596 morpholino (MO) and compared with control morpholino injected sibling controls.
597 Morpholino sequences used: control morpholino (5'-
598 CCTCTTACCTCAGTTACAATTTATA-3'), splice-blocking *sting* (5'-
599 GCCATGATACCTGGA-3') (van der Vaart et al., 2014), translation-blocking *cgas* (5'-
600 CTGGTCTCCTGTGGCTGCTCATGAT-3') (Ge et al., 2015), and splice-blocking *cgas* (5'-
601 CAGATTGTGGGCATAGTATACCTCT-3'). Translation-blocking morpholino was

validated by western blot, and splice-blocking morpholinos were validated by RT-PCR. For construct injections, *tol2* RNA was injected with either 100μg of *ubi:gfp* or 200 μg of *fli1:M27RNASEH1-GFP*. Embryos were scored at ~40 hpf for presence of the transgene according to GFP expression using a fluorescent dissecting scope.

Whole-mount *in situ* hybridization

In situ hybridization steps were performed as described previously by Thisse *et al.* (Thisse and Thisse, 2008). with minor modifications: before proteinase K permeabilization, embryos older than 28 hpf were bleached after re-hydration to remove pigmentation. The bleaching was done for 5-10 minutes using a bleaching solution of 0.8% KOH, 0.9% H_2O_2 and 0.1% Tween 20. Embryos were then scored manually, imaged and genotyped. The *runx1* (Burns et al., 2002), *flt4* (Lawson et al., 2001), and *notch1b* (Lawson et al., 2001) probes were used, and *runx1* levels were quantified using FIJI in a similar manner to previously described methods (Dobrzycki et al., 2018).

Flow cytometry

At 40 hpf, mutant and sibling embryos were binned based on morphological differences. For generation of single-cell suspensions, 10-20 embryos were first removed from their chorions using pronase (Roche), and then homogenized by manual dissociation using a sterile razor blade followed by digestion with Liberase (Roche). For the digestion, dissociated embryos were resuspended in 600μl 1× Dulbeccos-PBS (D-PBS) (Life Technologies) supplemented with a 1:65 dilution (9.23 μl) of 5 mg/ml Liberase and then incubated at 37°C for 7 minutes. The reaction was stopped with the addition of 5% (30μl) fetal bovine serum (FBS) (Life Technologies). The cells were then filtered through a 40-μm cell strainer (Falcon) and pelleted by centrifugation at 3000 rpm for 5 minutes. Cell pellets were resuspended in 400-700 μl FACS buffer (0.9× D-PBS, 5% FBS, 1% Penn/Strep (Life Technologies)). DAPI (4′,6-diamidino-2-phenylindole) was added to a final concentration of 1 μg/ml to facilitate exclusion of dead cells from the analysis. Samples were analyzed with a LSRII flow cytometer (BD Biosciences) and FlowJo software. Cells from non-fluorescent embryos were used to set gates above background. Quantification for the absolute number of cells was performed by acquiring all events in

633 a tube on the flow cytometer to determine the total number of target cells. This number
634 was then divided by the total number of embryos analyzed to calculate the number of
635 target cells per embryo.
636
637 **Cell cycle analysis**
638 For EdU incorporation experiments, 38 hpf embryos were binned into 6 well plates, with
639 ~30 embryos per well. Embryos were incubated with 20mM EdU for 2 hours using the
640 Click-IT EdU Flow Cytometry Assay Kit according to the manufacturer's instructions
641 (Invitrogen C10424). Generation of single-cell suspensions of 40 hpf embryos was
642 accomplished as described above. Samples were analyzed with a LSRII flow cytometer
643 (BD Biosciences) and FlowJo software.
644
645 **Single-cell immunofluorescence**
646 **Zebrafish**
647 Single-cell suspensions were prepared as described above, and cell staining was
648 performed as described in Sorrells & Nik *et al.* (Sorrells et al., 2018). The mouse
649 monoclonal S9.6 (Boguslawski et al., 1986) primary antibody was used at a dilution of
650 1:50, the pTBK1 primary antibody (Cell Signaling Technology) was used at a dilution of
651 1:100, and the cGAS antibody (Abcam) was used at 1:200. All antibodies were diluted in
652 blocking buffer (5% bovine serum albumin (BSA)/0.2% milk/PBS) and incubated for 3 h
653 at room temperature followed by 3 washes in 0.1% Tween 20/PBS. The goat anti-mouse
654 AlexaFluor-488 IgG2a (Thermo Fisher Scientific) and goat anti-mouse AlexaFluor-488
655 IgG (Fisher) secondary antibody were used at a dilution of 1:1000, while the mouse anti-
656 rabbit AlexaFluor-488 (Thermo Fisher Scientific) secondary antibody was used at a
657 dilution of 1:500; all secondary antibodies were incubated for 1 h at room temperature
658 followed by 3 washes in 0.1% Tween 20/PBS. After final washes, cells were mounted
659 with DAPI-Fluoromount-G (Southern Biotech), covered with a 25×25 mm glass coverslip
660 and sealed with clear nail polish. Fluorescence intensity measurements of S9.6 were
661 performed using FIJI. To quantify nuclear R-loop and cGAS levels, DAPI staining was
662 used to define the nuclear border, and only R-loops or cGAS within the nucleus were

663 measured using Fiji. To quantify non-nuclear R-loop levels, we measured R-loop levels
664 using Fiji that were adjacent to, but not within the DAPI stained regions.
665
666 **Human cells**
667 HEK-293 cells grown on 25×25 mm glass coverslips were fixed in 4% paraformaldehyde
668 (PFA) for 10 minutes followed by two 5-minute washes in D-PBS. Cells were
669 permeabilized with 0.1% Triton-X in D-PBS for 10 minutes followed by two 10-minute
670 washes in D-PBS. Cells were then blocked in 1% FBS/D-PBS for 1 h. The mouse
671 monoclonal S9.6 (Boguslawski et al., 1986) primary antibody was used at a dilution of
672 1:100 in blocking buffer for 2 h at room temperature followed by 5 washes in D-PBS for 5
673 minutes each. The goat anti-mouse AlexaFluor-488 IgG2a was used at a dilution of 1:500
674 and incubated for 1 h at room temperature followed by 5 washes in D-PBS for 5 minutes
675 each. After final washes, cells were mounted with DAPI-Fluoromount-G, placed on a
676 glass slide and sealed with clear nail polish. Fluorescence intensity measurements of
677 S9.6 were performed using FIJI.
678
679 **Fluorescent microscopy and time-lapse imaging**
680 Embryos were imaged using a Leica SP5 inverted confocal microscope (Leica). For time-
681 lapse imaging, embryos were anesthetized with tricaine and then embedded in agarose
682 (1% in E3 medium) containing tricaine in a zWEDGI device (Huemer et al., 2017). *z*-
683 stacks were taken every 20 minutes over a six-hour period. Movies were created using
684 Fiji software.
685
686 **Image acquisition and processing**
687 Brightfield and fluorescent images were acquired using a Zeiss Axio Discovery.V8 with
688 an AxioCam HRc camera. Confocal images for zebrafish single-cell immunofluorescence
689 were acquired using a Leica SP5 AOBS Inverted DMI6000 microscope with a 63X oil
690 objective and zoomed in fields were taken at 6.5X zoom. FIJI/ImageJ was used for
691 quantification of images (Collins, 2007). Flowjo, GraphPad Prism 8, Adobe Photoshop
692 and Adobe Illustrator were used to generate figures.
693

694 **RNA-sequencing and splicing analysis**

695 *cd41:*gfp^{+} HSPCs from *ddx41* mutants and siblings at 40 hpf were isolated by FACS on
696 a Mo-Flo cell sorter (Beckman Coulter). RNA from these cells was subsequently isolated
697 using the Zymo Quick-RNA microprep kit (R2080) according to the manufacturer
698 specification. The TURBO DNA-free kit (Life Technologies) was used for DNA removal
699 following RNA extraction. RNA and library quality was assessed using a bioanalyzer Pico
700 Chip (Agilent) by the Einstein Genomics core facility. Libraries for RNA samples with
701 RIN≥9 were prepared by BGI using the Nugene low-input RNA sequencing kit. Three
702 independent biological replicates were analyzed. On average, approximately 30 million
703 paired-end 150 base pair sequencing reads were acquired per sample using the Illumina
704 NovoSeq platform. Zebrafish RNA sequencing reads were first trimmed with TRIM
705 GALORE software to check for quality of 150bp reads and then mapped to the *D. rerio*
706 GRcz10 genome build using STAR Aligner (Dobin et al., 2013) (version 2.4.2a). The
707 GRCz10.85 gtf annotation file was used. DE-Seq2 (Love et al., 2014) was used via
708 DEApp (Li and Andrade, 2017) to determine differential gene expression (defined as 2-
709 fold change, FDR $p<0.05$) between *ddx41* mutant and sibling control HSPCs. Pathway
710 analysis was analyzed using the Molecular Signatures Database (MSigDB) (Liberzon et
711 al., 2015; Subramanian et al., 2005). Differential inclusion rates across 17,040 splicing
712 events between *ddx41* mutants and siblings were calculated using rMATS (Shen et al.,
713 2014) (replicate Multivariate Analysis of Transcript Splicing). Splicing events with
714 inclusion differences ≥ 10% and FDR ≤ 0.01 were considered significantly altered.
715 Pathway analysis for alternatively spliced genes was also performed using MSigDB.
716 RNA-seq data have been deposited at Gene Expression Omnibus (GEO) under
717 GSE146995.
718
719 **RT-PCR**
720 Mutant and sibling embryos were pooled based on morphology. Zebrafish embryo RNA
721 was isolated from pools of 20-40 embryos at the 40 hpf stage using the TRIzol reagent
722 (Thermo Fisher; 15596026) according to the manufacturer's protocol. Embryos were
723 dissociated in TRIzol using a pestle. The High Capacity cDNA reverse transcriptase kit
724 (Thermo Fisher; 4368814) was used to generate cDNA from 300ng input RNA. The cDNA

was diluted 1:50 in nuclease-free water. Data were averaged from at least three independent replicates per sample. qPCR primers are in Key Resources Table S6. For RT-PCR analysis for morpholino functionality, wild-type embryos were first injected with control, splice-blocking *sting*, or splice-blocking *cgas* morpholinos. At 24 hpf, zebrafish embryo RNA was isolated and used to generate cDNA similar to the method described above. The cDNA was diluted 1:5 in nuclease-free water. The cDNA was then amplified using the RT-PCR primers in Key Resources Table S6, and are run on an agarose gel to confirm proper exon-skipping induced by the splice-blocking morpholinos.

Western Analysis

Embryos were pooled for protein collection on the basis of phenotype. Cell lysates of deyolked embryos were prepared with RIPA buffer (Boston BioProducts). Proteins were resolved by SDS-PAGE and transferred to nitrocellulose membranes (Bio-Rad), followed by immunoblotting with primary and secondary antibodies. Primary antibodies were: mouse anti-TBK1 (Cell Signaling Technology, 1:1000), mouse anti-pTBK1 (Cell Signaling Technology, 1:1000), rabbit anti-cGAS (Abcam, 1:700) and rabbit anti-GAPDH (Abcam, 1:10000). Secondary antibodies were: goat anti-mouse IgG-HRP (Santa Cruz, 1:2000) and mouse anti-rabbit IgG-HRP (Santa Cruz, 1:2000 for rabbit anti-cGAS, and 1:5000 for rabbit anti-GAPDH). Catalog numbers for all antibodies are listed in the Key Resources Table S6.

Patient database and survival data

Gene expression data from 183 MDS $CD34^+$ samples and 17 controls were obtained from Gene Expression Omnibus (GEO) (GSE19429) (Pellagatti et al., 2010). Clinical annotations and overall survival were compared to gene expression as performed previously (Schinke et al., 2015). Pathway analysis between *DDX41* low and *DDX41* high patient $CD34^+$ cells was performed using MSigDB, with clinical annotation being performed as described above. Each cell represents the correlation between the 183 values of the clinical variable (column), and the 183 values of the gene probe (row) matched by patient.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics

Experiments were performed with a minimum of three replicates. Statistical analyses were performed as indicated in each figure using unpaired Student's t-test or a one-way ANOVA with Tukey's multiple testing correction as appropriate; error bars indicate the standard deviation of mean, unless otherwise indicated.

DATA AND CODE AVAILABILITY

The accession numbers for the RNA-Seq data reported in this paper are under Gene Expression Omnibus (GEO) GSE146995.

766 **REFERENCES**

767 Ablasser, A., Schmid-Burgk, J.L., Hemmerling, I., Horvath, G.L., Schmidt, T., Latz, E., and
768 Hornung, V. (2013). Cell intrinsic immunity spreads to bystander cells via the intercellular
769 transfer of cGAMP. Nature *503*, 530-534.
770 Bai, J., Cervantes, C., Liu, J., He, S., Zhou, H., Zhang, B., Cai, H., Yin, D., Hu, D., Li, Z., *et al.*
771 (2017). DsbA-L prevents obesity-induced inflammation and insulin resistance by suppressing
772 the mtDNA release-activated cGAS-cGAMP-STING pathway. Proc Natl Acad Sci U S A *114*,
773 12196-12201.
774 Bakhoum, S.F., Ngo, B., Laughney, A.M., Cavallo, J.A., Murphy, C.J., Ly, P., Shah, P., Sriram,
775 R.K., Watkins, T.B.K., Taunk, N.K., *et al.* (2018). Chromosomal instability drives metastasis
776 through a cytosolic DNA response. Nature *553*, 467-472.
777 Barreyro, L., Chlon, T.M., and Starczynowski, D.T. (2018). Chronic immune response
778 dysregulation in MDS pathogenesis. Blood *132*, 1553-1560.
779 Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y., and Traver, D. (2010).
780 Haematopoietic stem cells derive directly from aortic endothelium during development. Nature
781 *464*, 108-111.
782 Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L.,
783 and Carrico, R.J. (1986). Characterization of monoclonal antibody to DNA.RNA and its
784 application to immunodetection of hybrids. J Immunol Methods *89*, 123-130.
785 Boisset, J.C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C.
786 (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium.
787 Nature *464*, 116-120.
788 Burns, C.E., DeBlasio, T., Zhou, Y., Zhang, J., Zon, L., and Nimer, S.D. (2002). Isolation and
789 characterization of runxa and runxb, zebrafish members of the runt family of transcriptional
790 regulators. Experimental hematology *30*, 1381-1389.
791 Chen, L., Chen, J.Y., Huang, Y.J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H.,
792 *et al.* (2018). The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes
793 Induced by High-Risk Splicing Factor Mutations. Mol Cell *69*, 412-425 e416.
794 Collins, T.J. (2007). ImageJ for microscopy. Biotechniques *43*, 25-30.
795 Crossley, M.P., Bocek, M., and Cimprich, K.A. (2019). R-Loops as Cellular Regulators and
796 Genomic Threats. Mol Cell *73*, 398-411.
797 Curran, E., Chen, X., Corrales, L., Kline, D.E., Dubensky, T.W., Jr., Duttagupta, P., Kortylewski,
798 M., and Kline, J. (2016). STING Pathway Activation Stimulates Potent Immunity against Acute
799 Myeloid Leukemia. Cell reports *15*, 2357-2366.

800 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
801 M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*,
802 15-21.
803 Dobrzycki, T., Krecsmarik, M., Bonkhofer, F., Patient, R., and Monteiro, R. (2018). An optimised
804 pipeline for parallel image-based quantification of gene expression and genotyping after in situ
805 hybridisation. Biol Open *7*.
806 Espin-Palazon, R., Stachura, D.L., Campbell, C.A., Garcia-Moreno, D., Del Cid, N., Kim, A.D.,
807 Candel, S., Meseguer, J., Mulero, V., and Traver, D. (2014). Proinflammatory signaling
808 regulates hematopoietic stem cell emergence. Cell *159*, 1070-1085.
809 Espin-Palazon, R., Weijts, B., Mulero, V., and Traver, D. (2018). Proinflammatory Signals as
810 Fuel for the Fire of Hematopoietic Stem Cell Emergence. Trends Cell Biol *28*, 58-66.
811 Fagerberg, L., Hallstrom, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka,
812 M., Tahmasebpoor, S., Danielsson, A., Edlund, K., *et al.* (2014). Analysis of the human tissue-
813 specific expression by genome-wide integration of transcriptomics and antibody-based
814 proteomics. Mol Cell Proteomics *13*, 397-406.
815 Fang, J., Muto, T., Kleppe, M., Bolanos, L.C., Hueneman, K.M., Walker, C.S., Sampson, L.,
816 Wellendorf, A.M., Chetal, K., Choi, K., *et al.* (2018). TRAF6 Mediates Basal Activation of NF-
817 kappaB Necessary for Hematopoietic Stem Cell Homeostasis. Cell Rep *22*, 1250-1262.
818 Ge, R., Zhou, Y., Peng, R., Wang, R., Li, M., Zhang, Y., Zheng, C., and Wang, C. (2015).
819 Conservation of the STING-Mediated Cytosolic DNA Sensing Pathway in Zebrafish. J Virol *89*,
820 7696-7706.
821 Gentili, M., Lahaye, X., Nadalin, F., Nader, G.P.F., Lombardi, E.P., Herve, S., De Silva, N.S.,
822 Rookhuizen, D.C., Zueva, E., Goudot, C., *et al.* (2019). The N-Terminal Domain of cGAS
823 Determines Preferential Association with Centromeric DNA and Innate Immune Activation in the
824 Nucleus. Cell reports *26*, 3798.
825 Gunther, C., Kind, B., Reijns, M.A., Berndt, N., Martinez-Bueno, M., Wolf, C., Tungler, V.,
826 Chara, O., Lee, Y.A., Hubner, N., *et al.* (2015). Defective removal of ribonucleotides from DNA
827 promotes systemic autoimmunity. J Clin Invest *125*, 413-424.
828 Haferlach, T., Nagata, Y., Grossmann, V., Okuno, Y., Bacher, U., Nagae, G., Schnittger, S.,
829 Sanada, M., Kon, A., Alpermann, T., *et al.* (2014). Landscape of genetic lesions in 944 patients
830 with myelodysplastic syndromes. Leukemia *28*, 241-247.
831 He, Q., Zhang, C., Wang, L., Zhang, P., Ma, D., Lv, J., and Liu, F. (2015). Inflammatory
832 signaling regulates hematopoietic stem and progenitor cell emergence in vertebrates. Blood
833 *125*, 1098-1106.

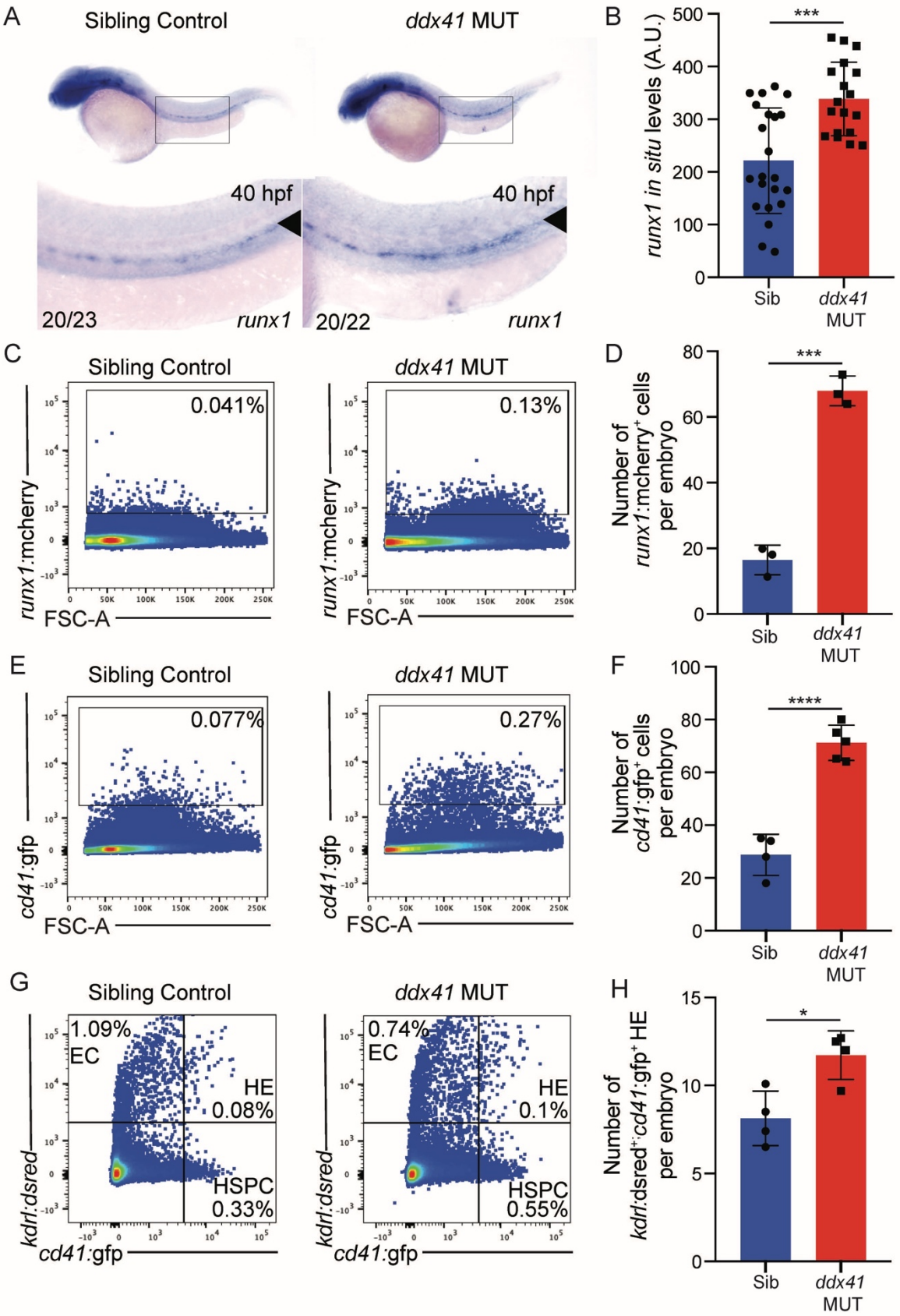
834 Jiang, H., Xue, X., Panda, S., Kawale, A., Hooy, R.M., Liang, F., Sohn, J., Sung, P., and
835 Gekara, N.O. (2019). Chromatin-bound cGAS is an inhibitor of DNA repair and hence
836 accelerates genome destabilization and cell death. EMBO J *38*, e102718.
837 Kanther, M., Sun, X., Muhlbauer, M., Mackey, L.C., Flynn, E.J., 3rd, Bagnat, M., Jobin, C., and
838 Rawls, J.F. (2011). Microbial colonization induces dynamic temporal and spatial patterns of NF-
839 kappaB activation in the zebrafish digestive tract. Gastroenterology *141*, 197-207.
840 Kettleborough, R.N., Busch-Nentwich, E.M., Harvey, S.A., Dooley, C.M., de Bruijn, E., van
841 Eeden, F., Sealy, I., White, R.J., Herd, C., Nijman, I.J., *et al.* (2013). A systematic genome-wide
842 analysis of zebrafish protein-coding gene function. Nature *496*, 494-497.
843 Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel
844 type of cell transition. Nature *464*, 112-115.
845 Lawrence, C. (2011). Advances in zebrafish husbandry and management. Methods Cell Biol
846 *104*, 429-451.
847 Lawson, N.D., Scheer, N., Pham, V.N., Kim, C.H., Chitnis, A.B., Campos-Ortega, J.A., and
848 Weinstein, B.M. (2001). Notch signaling is required for arterial-venous differentiation during
849 embryonic vascular development. Development *128*, 3675-3683.
850 Lawson, N.D., and Weinstein, B.M. (2002). In vivo imaging of embryonic vascular development
851 using transgenic zebrafish. Dev Biol *248*, 307-318.
852 Lemos, H., Mohamed, E., Huang, L., Ou, R., Pacholczyk, G., Arbab, A.S., Munn, D., and Mellor,
853 A.L. (2016). STING Promotes the Growth of Tumors Characterized by Low Antigenicity via IDO
854 Activation. Cancer Res *76*, 2076-2081.
855 Li, X., Shu, C., Yi, G., Chaton, C.T., Shelton, C.L., Diao, J., Zuo, X., Kao, C.C., Herr, A.B., and
856 Li, P. (2013). Cyclic GMP-AMP synthase is activated by double-stranded DNA-induced
857 oligomerization. Immunity *39*, 1019-1031.
858 Li, Y., and Andrade, J. (2017). DEApp: an interactive web interface for differential expression
859 analysis of next generation sequence data. Source Code Biol Med *12*, 2.
860 Li, Y., Esain, V., Teng, L., Xu, J., Kwan, W., Frost, I.M., Yzaguirre, A.D., Cai, X., Cortes, M.,
861 Maijenburg, M.W., *et al.* (2014). Inflammatory signaling regulates embryonic hematopoietic stem
862 and progenitor cell production. Genes Dev *28*, 2597-2612.
863 Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015).
864 The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-
865 425.
866 Lim, Y.W., Sanz, L.A., Xu, X., Hartono, S.R., and Chedin, F. (2015). Genome-wide DNA
867 hypomethylation and RNA:DNA hybrid accumulation in Aicardi-Goutieres syndrome. Elife *4*.

868 Lin, H.F., Traver, D., Zhu, H., Dooley, K., Paw, B.H., Zon, L.I., and Handin, R.I. (2005). Analysis
869 of thrombocyte development in CD41-GFP transgenic zebrafish. Blood *106*, 3803-3810.
870 Lindsley, R.C., and Ebert, B.L. (2013). Molecular pathophysiology of myelodysplastic
871 syndromes. Annu Rev Pathol *8*, 21-47.
872 Liu, H., Zhang, H., Wu, X., Ma, D., Wu, J., Wang, L., Jiang, Y., Fei, Y., Zhu, C., Tan, R., *et al.*
873 (2018). Nuclear cGAS suppresses DNA repair and promotes tumorigenesis. Nature *563*, 131-
874 136.
875 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
876 dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
877 Mankan, A.K., Schmidt, T., Chauhan, D., Goldeck, M., Honing, K., Gaidt, M., Kubarenko, A.V.,
878 Andreeva, L., Hopfner, K.P., and Hornung, V. (2014). Cytosolic RNA:DNA hybrids activate the
879 cGAS-STING axis. EMBO J *33*, 2937-2946.
880 Meeker, N.D., Hutchinson, S.A., Ho, L., and Trede, N.S. (2007). Method for isolation of PCR-
881 ready genomic DNA from zebrafish tissues. Biotechniques *43*, 610, 612, 614.
882 Montoro, J., Gallur, L., Merchan, B., Molero, A., Roldan, E., Martinez-Valle, F., Villacampa, G.,
883 Navarrete, M., Ortega, M., Castellvi, J., *et al.* (2018). Autoimmune disorders are common in
884 myelodysplastic syndrome patients and confer an adverse impact on outcomes. Ann Hematol
885 *97*, 1349-1356.
886 Nakayama, S., Ikenaga, T., Kawakami, K., Ono, F., and Hatta, K. (2012). Transgenic line with
887 gal4 insertion useful to study morphogenesis of craniofacial perichondrium, vascular
888 endothelium-associated cells, floor plate, and dorsal midline radial glia during zebrafish
889 development. Dev Growth Differ *54*, 202-215.
890 Ng, K.W., Marshall, E.A., Bell, J.C., and Lam, W.L. (2018). cGAS-STING and Cancer:
891 Dichotomous Roles in Tumor Immunity and Development. Trends Immunol *39*, 44-54.
892 Nguyen, H.D., Leong, W.Y., Li, W., Reddy, P.N.G., Sullivan, J.D., Walter, M.J., Zou, L., and
893 Graubert, T.A. (2018). Spliceosome Mutations Induce R Loop-Associated Sensitivity to ATR
894 Inhibition in Myelodysplastic Syndromes. Cancer Res *78*, 5363-5374.
895 Nik, S., Weinreb, J.T., and Bowman, T.V. (2017). Developmental HSC Microenvironments:
896 Lessons from Zebrafish. Adv Exp Med Biol *1041*, 33-53.
897 Ottersbach, K. (2019). Endothelial-to-haematopoietic transition: an update on the process of
898 making blood. Biochem Soc Trans *47*, 591-601.
899 Parvatiyar, K., Zhang, Z., Teles, R.M., Ouyang, S., Jiang, Y., Iyer, S.S., Zaver, S.A., Schenk,
900 M., Zeng, S., Zhong, W., *et al.* (2012). The helicase DDX41 recognizes the bacterial secondary

901 messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response.
902 Nat Immunol *13*, 1155-1161.
903 Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.C., Guan, A., Hesley, J.A., Miller,
904 S.C., Cromwell, E.F., Solow-Cordero, D.E., *et al.* (2009). A genome-wide siRNA screen reveals
905 diverse cellular processes and pathways that mediate genome stability. Mol Cell *35*, 228-239.
906 Pellagatti, A., Cazzola, M., Giagounidis, A., Perry, J., Malcovati, L., Della Porta, M.G.,
907 Jadersten, M., Killick, S., Verma, A., Norbury, C.J., *et al.* (2010). Deregulated gene expression
908 pathways in myelodysplastic syndrome hematopoietic stem cells. Leukemia *24*, 756-764.
909 Pietras, E.M. (2017). Inflammation: a key regulator of hematopoietic stem cell fate in health and
910 disease. Blood *130*, 1693-1698.
911 Polprasert, C., Schulze, I., Sekeres, M.A., Makishima, H., Przychodzen, B., Hosono, N., Singh,
912 J., Padgett, R.A., Gu, X., Phillips, J.G., *et al.* (2015). Inherited and Somatic Defects in DDX41 in
913 Myeloid Neoplasms. Cancer Cell *27*, 658-670.
914 Rajshekar, S., Yao, J., Arnold, P.K., Payne, S.G., Zhang, Y., Bowman, T.V., Schmitz, R.J.,
915 Edwards, J.R., and Goll, M. (2018). Pericentromeric hypomethylation elicits an interferon
916 response in an animal model of ICF syndrome. Elife *7*.
917 Sallman, D.A., and List, A. (2019). The central role of inflammatory signaling in the
918 pathogenesis of myelodysplastic syndromes. Blood *133*, 1039-1048.
919 Sawamiphak, S., Kontarakis, Z., and Stainier, D.Y. (2014). Interferon gamma signaling
920 positively regulates hematopoietic stem cell emergence. Developmental cell *31*, 640-653.
921 Schinke, C., Giricz, O., Li, W., Shastri, A., Gordon, S., Barreyro, L., Bhagat, T., Bhattacharyya,
922 S., Ramachandra, N., Bartenstein, M., *et al.* (2015). IL8-CXCR2 pathway inhibition as a
923 therapeutic strategy against MDS and AML stem cells. Blood *125*, 3144-3152.
924 Schneider, R.K., Schenone, M., Ferreira, M.V., Kramann, R., Joyce, C.E., Hartigan, C., Beier,
925 F., Brummendorf, T.H., Germing, U., Platzbecker, U., *et al.* (2016). Rps14 haploinsufficiency
926 causes a block in erythroid differentiation mediated by S100A8 and S100A9. Nat Med *22*, 288-
927 297.
928 Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014).
929 rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq
930 data. Proceedings of the National Academy of Sciences of the United States of America *111*,
931 E5593-5601.
932 Sorrells, S., Nik, S., Casey, M., Cameron, R.C., Truong, H., Toruno, C., Gulfo, M., Lowe, A.,
933 Jette, C., Stewart, R.A., and Bowman, T.V. (2018). Spliceosomal components protect
934 embryonic neurons from R-loop-mediated DNA damage and apoptosis. Dis Model Mech.

935 Sperling, A.S., Gibson, C.J., and Ebert, B.L. (2017). The genetics of myelodysplastic syndrome:
936 from clonal haematopoiesis to secondary leukaemia. Nat Rev Cancer *17*, 5-19.
937 Starczynowski, D.T., Kuchenbauer, F., Argiropoulos, B., Sung, S., Morin, R., Muranyi, A., Hirst,
938 M., Hogge, D., Marra, M., Wells, R.A., *et al.* (2010). Identification of miR-145 and miR-146a as
939 mediators of the 5q- syndrome phenotype. Nat Med *16*, 49-58.
940 Stavrou, S., Aguilera, A.N., Blouch, K., and Ross, S.R. (2018). DDX41 Recognizes RNA/DNA
941 Retroviral Reverse Transcripts and Is Critical for In Vivo Control of Murine Leukemia Virus
942 Infection. MBio *9*.
943 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,
944 Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set
945 enrichment analysis: a knowledge-based approach for interpreting genome-wide expression
946 profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.
947 Suzuki, Y., Holmes, J.B., Cerritelli, S.M., Sakhuja, K., Minczuk, M., Holt, I.J., and Crouch, R.J.
948 (2010). An upstream open reading frame and the context of the two AUG codons affect the
949 abundance of mitochondrial and nuclear RNase H1. Mol Cell Biol *30*, 5123-5134.
950 Tamplin, O.J., Durand, E.M., Carr, L.A., Childs, S.J., Hagedorn, E.J., Li, P., Yzaguirre, A.D.,
951 Speck, N.A., and Zon, L.I. (2015). Hematopoietic stem cell arrival triggers dynamic remodeling
952 of the perivascular niche. Cell *160*, 241-252.
953 Thisse, C., and Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish
954 embryos. Nat Protoc *3*, 59-69.
955 van der Vaart, M., Korbee, C.J., Lamers, G.E., Tengeler, A.C., Hosseini, R., Haks, M.C.,
956 Ottenhoff, T.H., Spaink, H.P., and Meijer, A.H. (2014). The DNA damage-regulated autophagy
957 modulator DRAM1 links mycobacterial recognition via TLR-MYD88 to autophagic defense
958 [corrected]. Cell Host Microbe *15*, 753-767.
959 Varney, M.E., Niederkorn, M., Konno, H., Matsumura, T., Gohda, J., Yoshida, N., Akiyama, T.,
960 Christie, S., Fang, J., Miller, D., *et al.* (2015). Loss of Tifab, a del(5q) MDS gene, alters
961 hematopoiesis through derepression of Toll-like receptor-TRAF6 signaling. J Exp Med *212*,
962 1967-1985.
963 Volkman, H.E., Cambier, S., Gray, E.E., and Stetson, D.B. (2019). Tight nuclear tethering of
964 cGAS is essential for preventing autoreactivity. Elife *8*.
965 Wang, I.X., Grunseich, C., Fox, J., Burdick, J., Zhu, Z., Ravazian, N., Hafner, M., and Cheung,
966 V.G. (2018). Human proteins that interact with RNA/DNA hybrids. Genome Res *28*, 1405-1414.

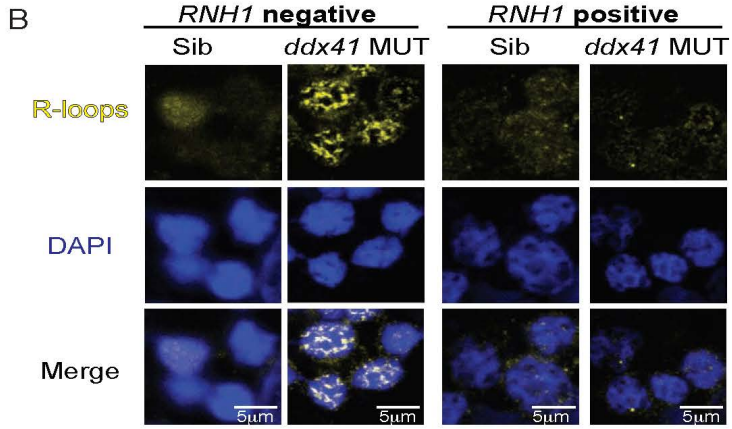
967 Yoneyama-Hirozane, M., Kondo, M., Matsumoto, S.I., Morikawa-Oki, A., Morishita, D.,
968 Nakanishi, A., Kawamoto, T., and Nakayama, M. (2017). High-Throughput Screening to Identify
969 Inhibitors of DEAD Box Helicase DDX41. SLAS Discov, 2472555217705952.
970 Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-
971 Otsubo, A., Kon, A., Nagasaki, M., *et al.* (2011). Frequent pathway mutations of splicing
972 machinery in myelodysplasia. Nature *478*, 64-69.
973 Zhang, Z., Bao, M., Lu, N., Weng, L., Yuan, B., and Liu, Y.J. (2013). The E3 ubiquitin ligase
974 TRIM21 negatively regulates the innate immune response to intracellular double-stranded DNA.
975 Nat Immunol *14*, 172-178.
976 Zhang, Z., Yuan, B., Bao, M., Lu, N., Kim, T., and Liu, Y.J. (2011). The helicase DDX41 senses
977 intracellular DNA mediated by the adaptor STING in dendritic cells. Nat Immunol *12*, 959-965.
978

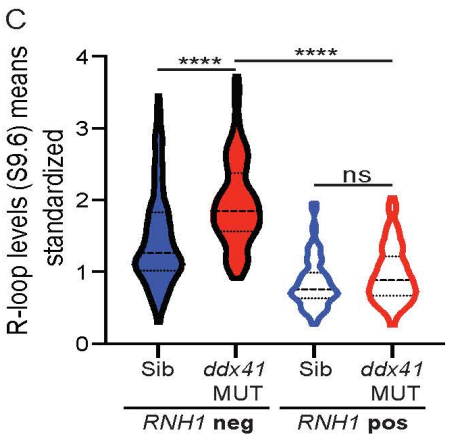


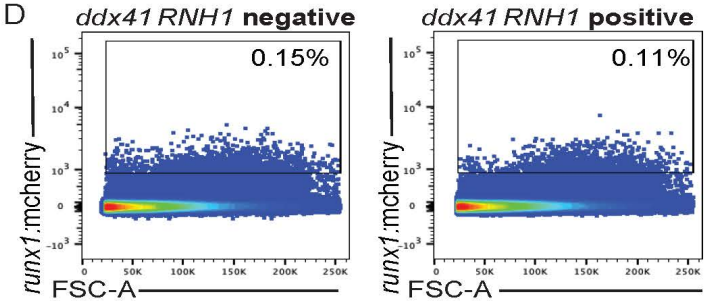
Weinreb et al. Figure 1

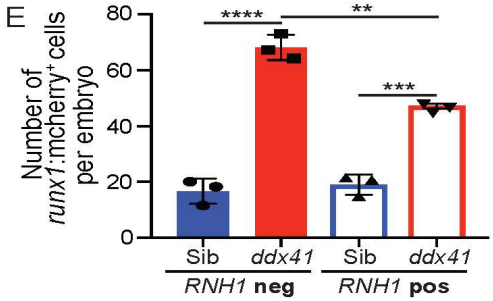
Figure 2

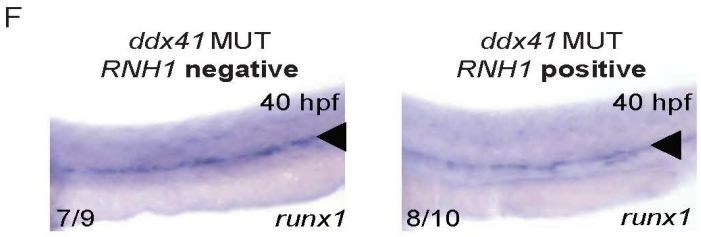


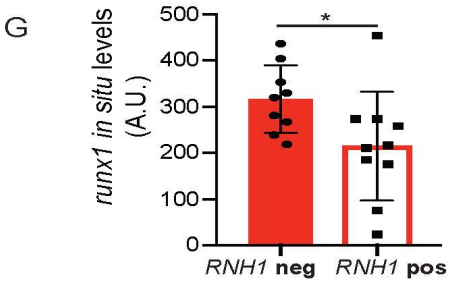


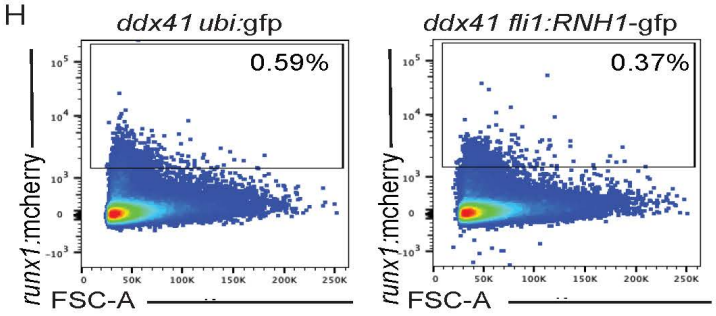


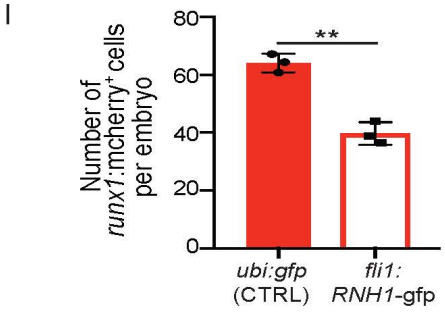






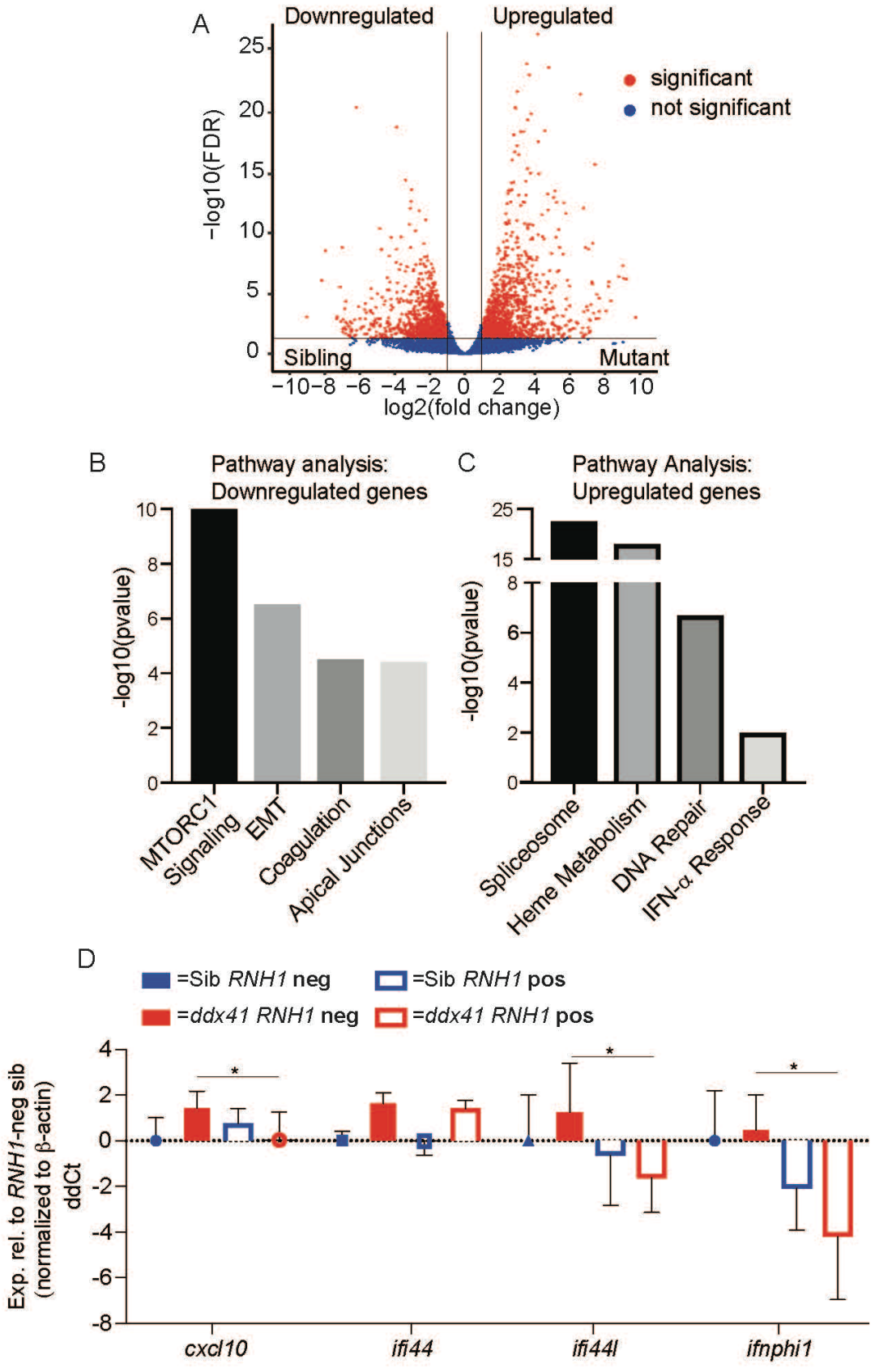






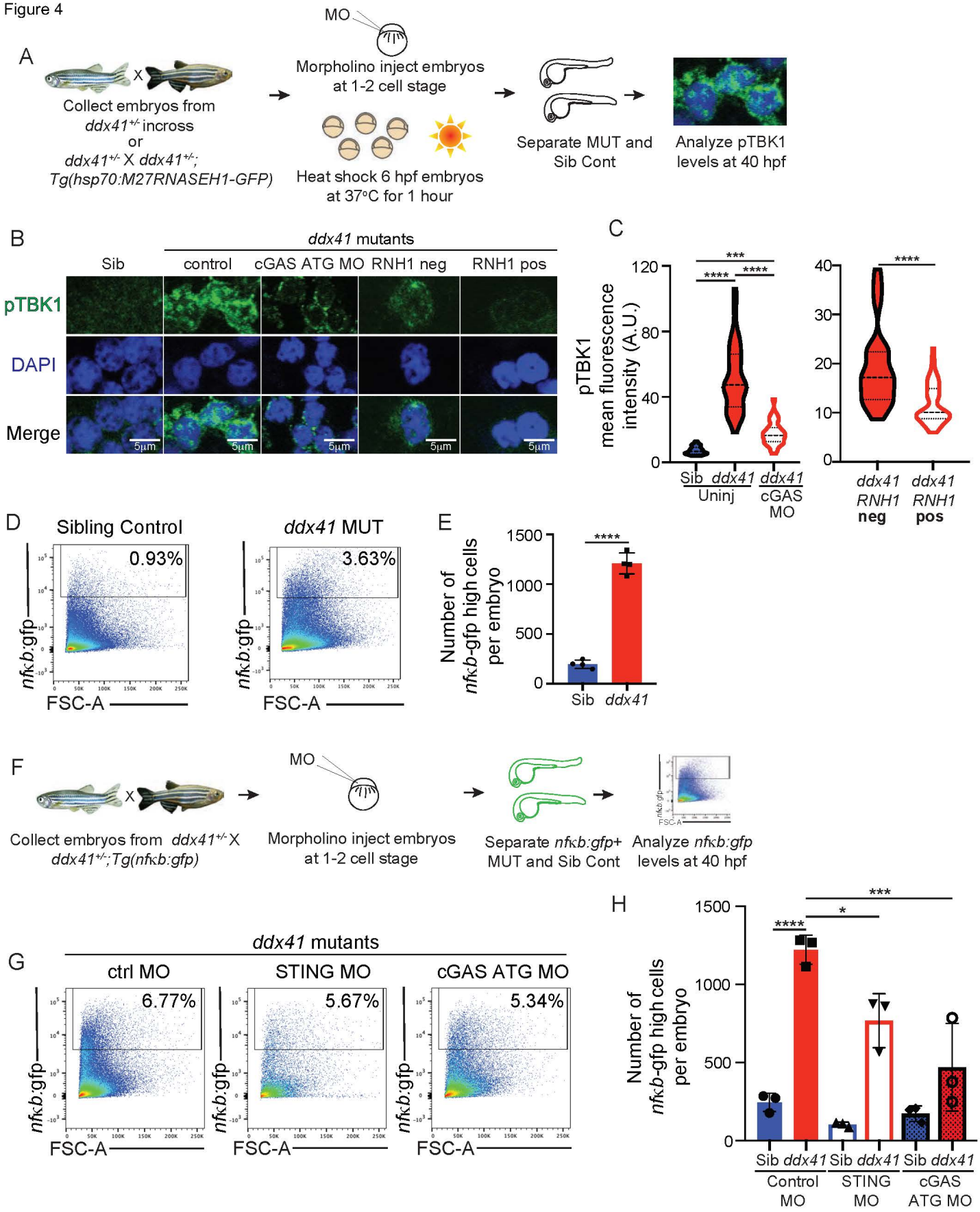
Weinreb et al. Figure 2

Figure 3



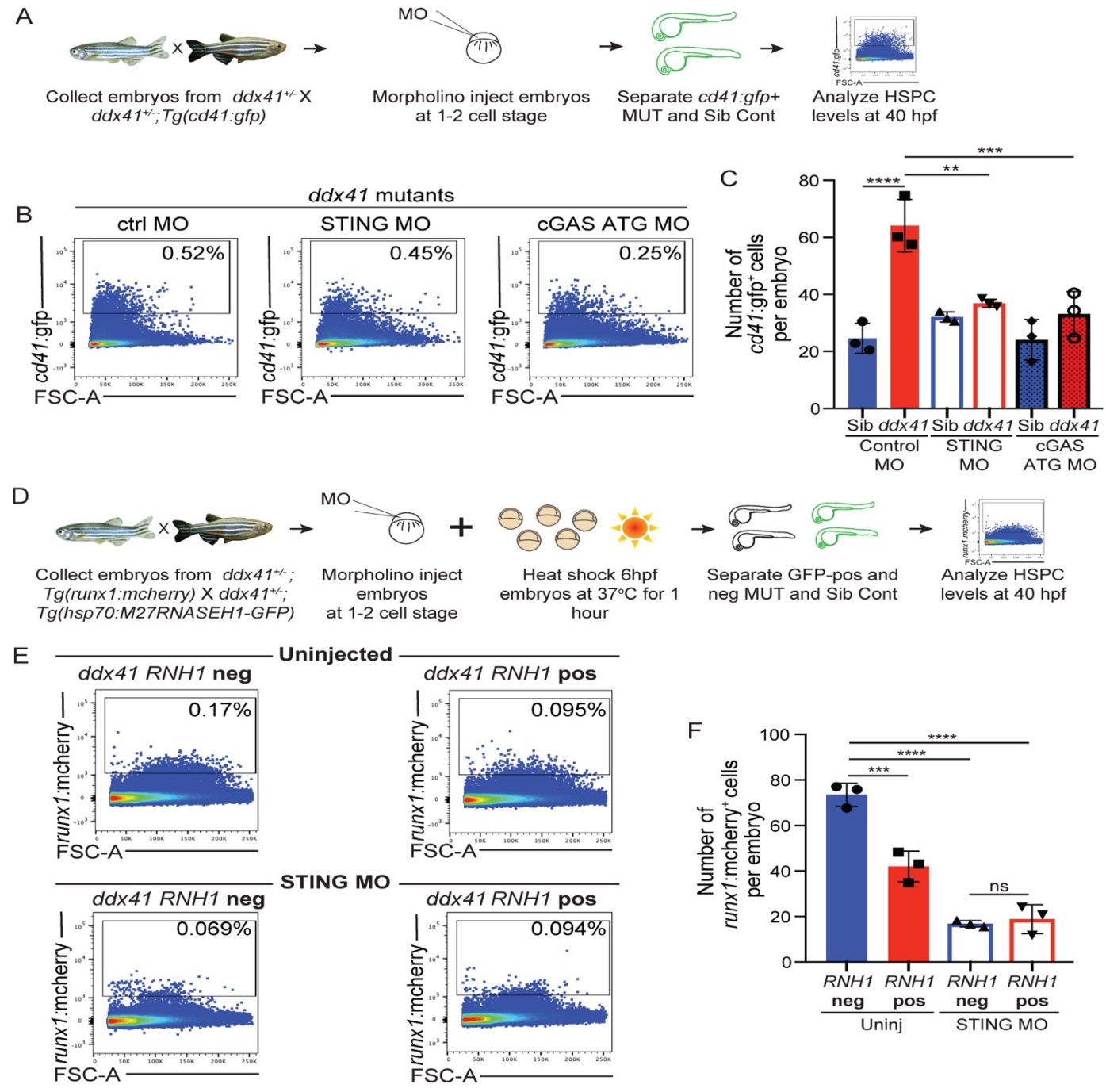
Weinreb et al. Figure 3

Figure 4



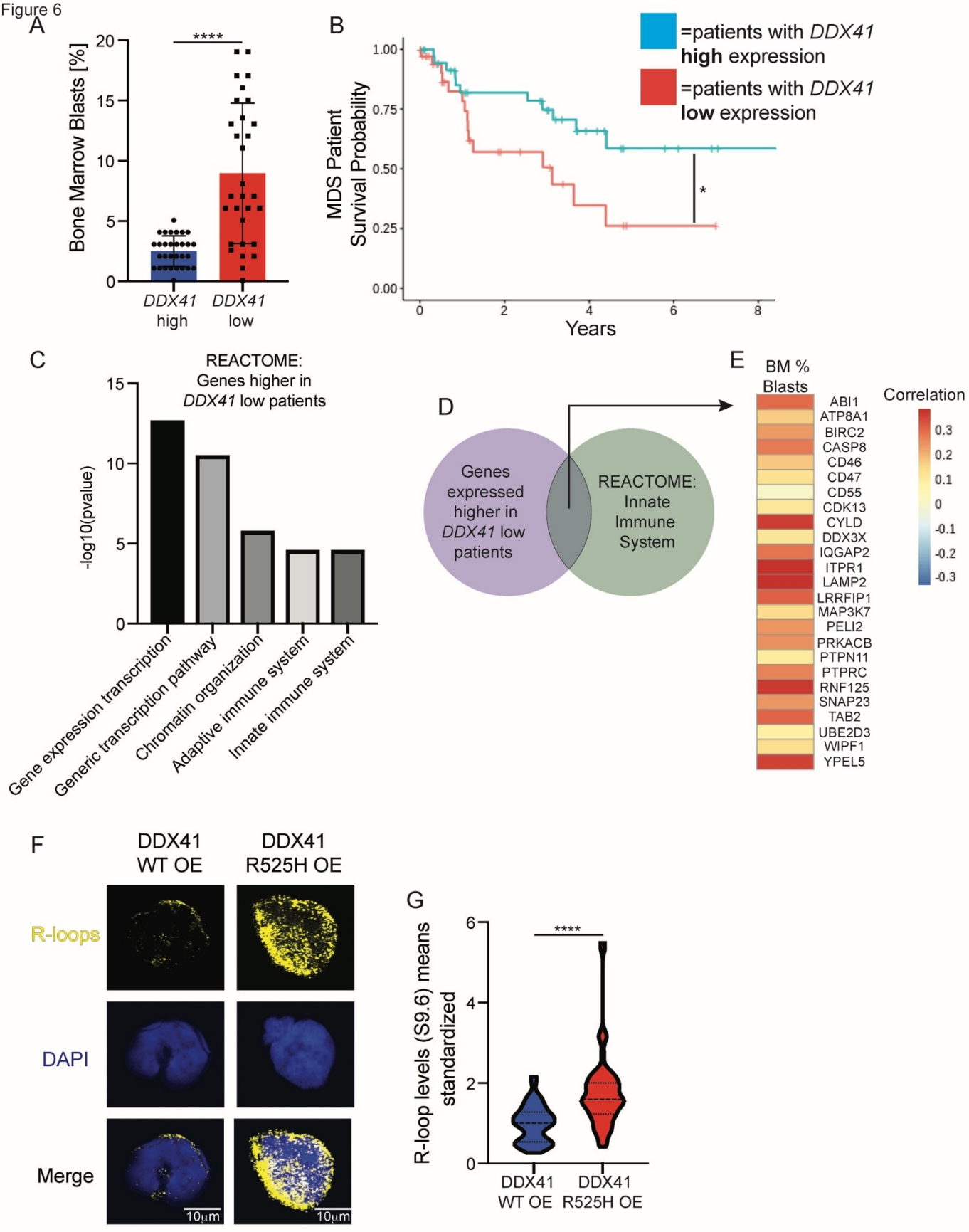
Weinreb et al. Figure 4

Figure 5



Weinreb et al. Figure 5

Figure 6



Weinreb et al. Figure 6

Click here to access/download;Supplemental Text and Figures;3rd submission Supplementary Figures Legend_v2.pdf

Supplemental Figures and Legends

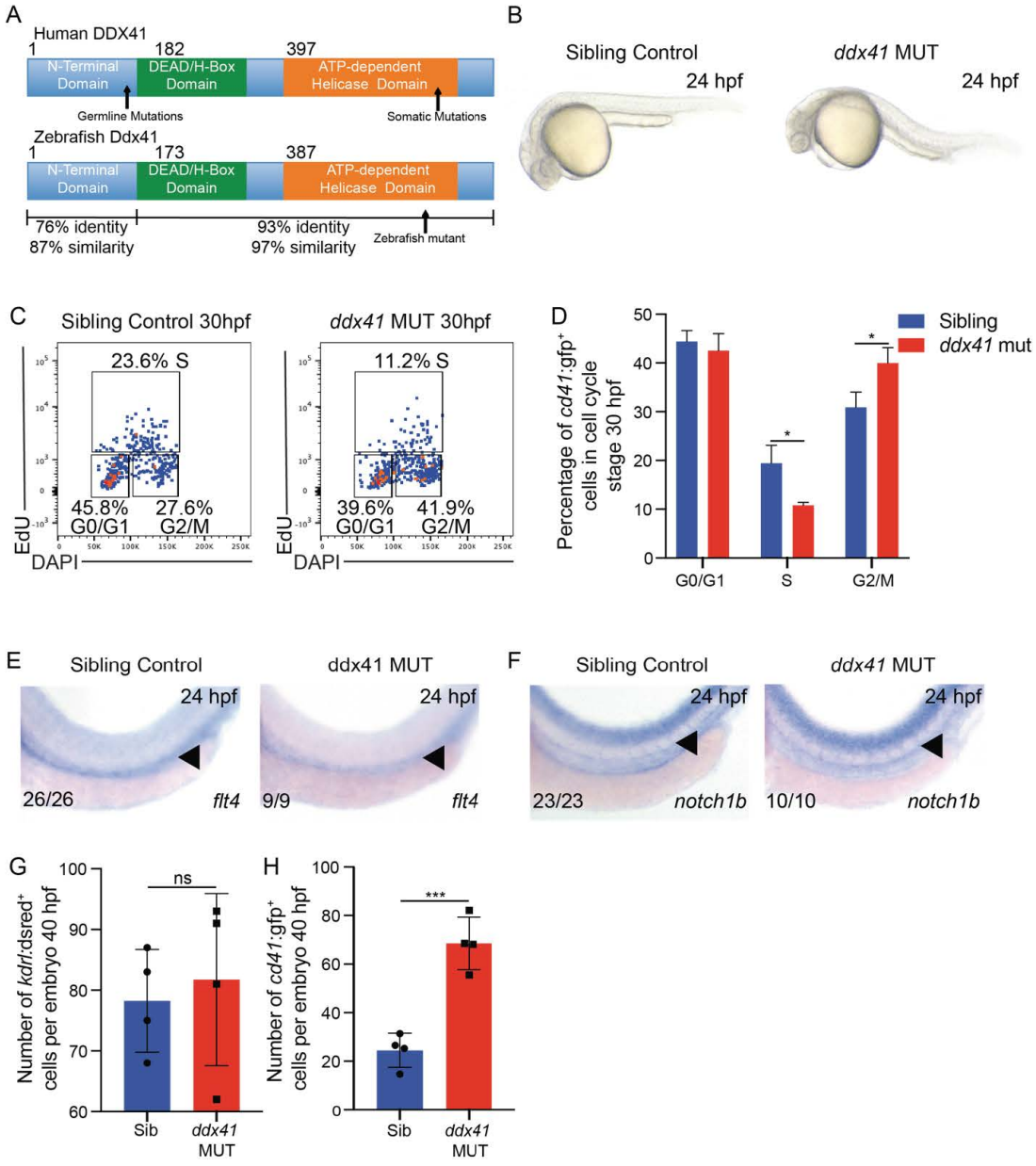


Figure S1. Ddx41 regulates HSPC number. A. Schema comparing sequence conservation between human and zebrafish DDX41 proteins, including the three major domains of the protein and the hotspot locations of germline and somatic mutations in *DDX41* from MDS/AML patient samples (which occur on separate alleles) and the

zebrafish mutation used in this study. **B**. Brightfield images of sibling controls (left) and *ddx41* mutants (right) at 24 hpf showing morphological differences, such as body curvature and mild brain death. **C.** Flow cytometry analysis of EdU incorporation in *cd41:*gfp^{+} HSPCs in sibling controls (left) and *ddx41* mutants (right) after a 2-hour pulse of EdU at 38 hpf. EdU incorporation (y-axis) and DAPI content (x-axis) were measured by flow cytometry at 40 hpf. **D.** Quantification of the percentage of cells in each cell cycle phase from (C). **E,F.** *In situ* hybridization of the vasculature markers *flt4* (C, vein) and *notch1b* (D, artery) at 40 hpf in sibling controls (left) and *ddx41* mutants (right). Numbers on bottom left corner indicate the fraction of embryos with the same phenotype as the one depicted in the image. Arrow marks the indicated regions. **G,H.** Graphs depicting the absolute number of *kdrl*:dsred^{+} endothelial cells (G) and *cd41:*gfp^{+} HSPCs (H) per embryo at 40hpf. Corresponding flow plots are shown in Figure 1G. Graphs display means ± standard deviations (stds) with p-values calculated with unpaired Student's t-test, *=p<0.05, ***=p<0.001, ns=not significant (p>0.05). N=3-5 replicates per experiment.

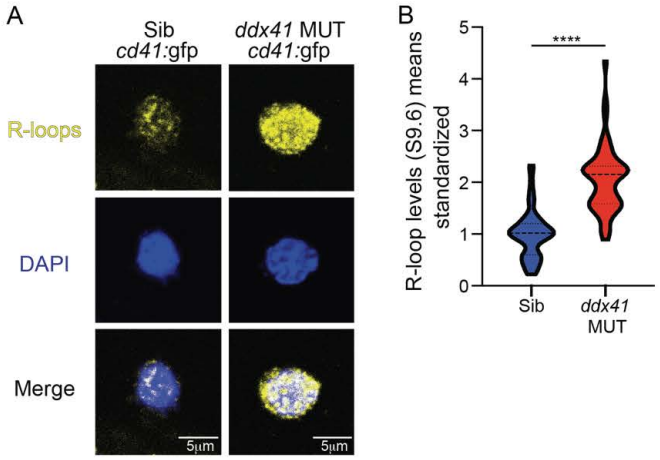


Figure S2. The control of R-loop levels by Ddx41 is critical for HSPC homeostasis. **A.** Confocal images showing immunofluorescence of R-loops and nuclei (DAPI) in sorted *cd41:*gfp^{+} HSPCs isolated from 40 hpf siblings (left) and *ddx41* mutants (right). Scale bar=5µm. **B**. Quantification of R-loop levels from (A). Graph displays mean ± standard deviation (std) with p-value calculated with unpaired Student's t-test, ****=p<0.0001. N=3 replicates per experiment.

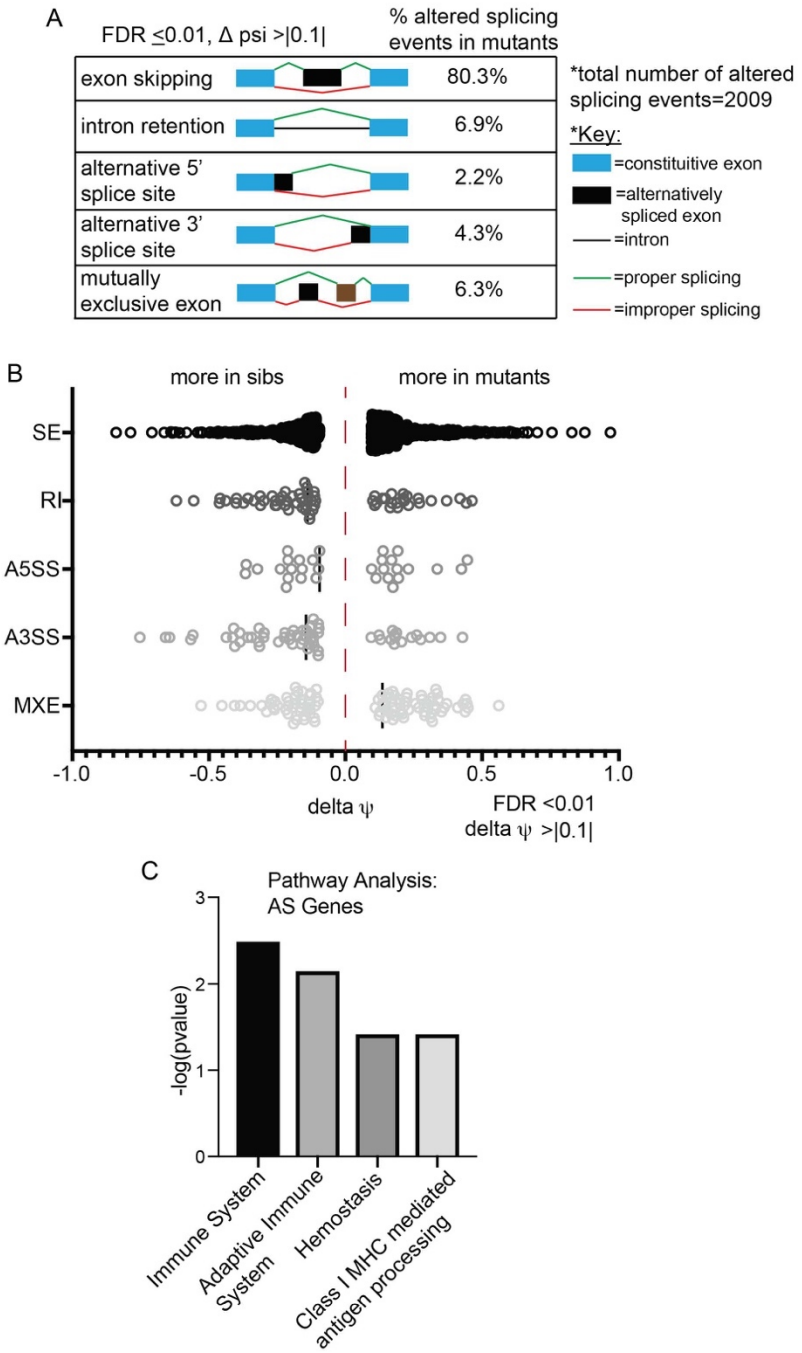


Figure S3. Splicing of immune system genes is altered in *ddx41* mutant HSPCs.

A. Schematic of the different splicing events detected by analysis with rMATS. Significant differences are defined as FDR ≤0.01 and Δ ψ >|0.1|. The percentage of alternatively

spliced events in mutants is based on the total number of altered splicing events in *ddx41* mutants as compared to siblings. Three replicates for both *ddx41* mutants and siblings were used to generate RNA-sequencing data. **B**. Graph depicting the $\Delta \psi$ of individual splicing events between sibling controls and *ddx41* mutants. **C**. Representative charts of pathways significantly enriched in alternatively spliced genes in *ddx41* mutant HSPCs compared to sibling controls as determined by MSigDB analysis.

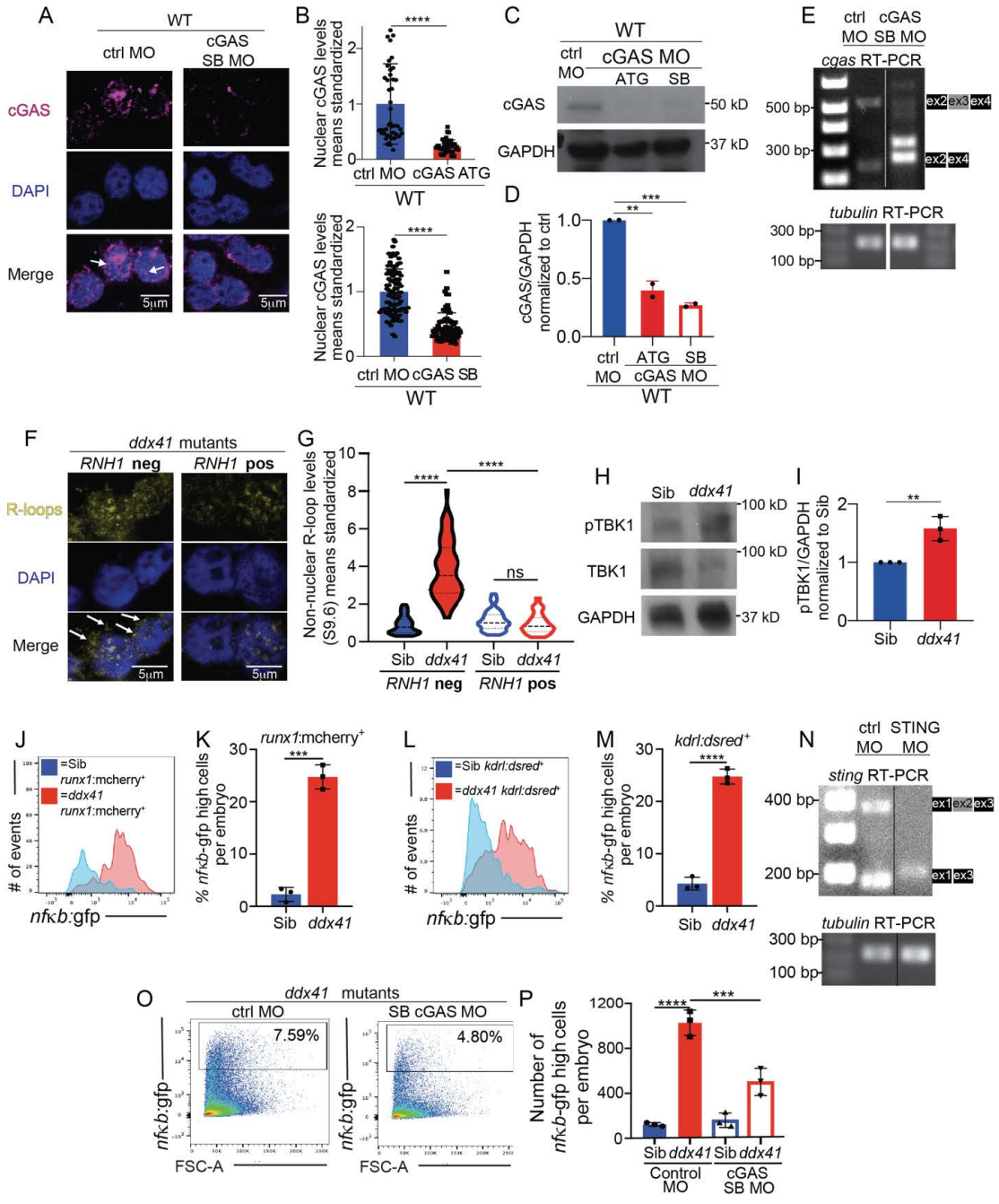


Figure S4. Ddx41 regulates inflammatory signaling via the cGAS-STING pathway.

A. Confocal images showing immunofluorescence of cGAS and nuclei (DAPI) in cells isolated from 28 hpf wildtype embryos injected with a control morpholino (left) and

wildtype embryos injected with a *cgas* splice-blocking (SB) morpholino (right). Scale bar=5μm. Arrows denote nuclear staining. **B.** Graph depicting quantification of nuclear cGAS levels in cells from wildtype embryos injected with control, *cgas* translation-blocking (ATG) (top), and *cgas* SB morpholinos (bottom). **C.** Western blots of cGAS and GAPDH (loading control) from wildtype embryos injected with a control (left), *cgas* ATG (middle), and *cgas* SB morpholinos (right) at 28 hpf. **D.** Graph depicting quantification of cGAS levels based on ratio of cGAS/GAPDH in (C). **E.** RT-PCR validation that the splice-blocking *cgas* morpholino induced exon 3 skipping (top). RT-PCR for *tubulin* as housekeeping control (bottom). **F.** Confocal images showing immunofluorescence of R-loops (RNA:DNA hybrid moiety) and nuclei (DAPI) in cells isolated from 28 hpf *ddx41* mutants that are either Tg*(hsp70:M27RNASEH1-GFP)*-negative (left) or *Tg(hsp70:M27RNASEH1-GFP)*-positive (right). Arrows indicate R-loops that are outside of nuclei (DAPI). Scale bar=5μm. **G.** Graph depicting quantification of non-nuclear R-loops in cells from siblings and *ddx41* mutants that are RNASEH1-GFP-negative or RNASEH1-GFP-positive. **H.** Western blots of Ser172-phosphorylated TBK1 (top), pan-TBK1 (middle), and GAPDH (bottom, loading control) from sibling controls (left) and *ddx41* mutants (right) at 40 hpf. **I.** Graph depicting quantification of pTBK1 levels based on ratio of pTBK1/GAPDH in (H). **J,L**. Flow cytometry analysis of *nfκb:gfp* levels in *runx1*:mcherry^{+} HSPCs (J) and *kdrl*:dsred^{+} endothelial cells (L) from sibling controls (blue) and *ddx41* mutants (red) at 40hpf. **K,M**. Graphs depicting percentage (%) of *nfκb:gfp*-high cells in *runx1*:mcherry^{+} HSPCs (K) and *kdrl*:dsred^{+} endothelial cells (M) at 40 hpf. **N.** RT-PCR validation that the splice-blocking *sting* morpholino induced exon 2 skipping (top). RT-PCR for *tubulin* as housekeeping control (bottom). **O.** Flow cytometry analysis of *nfκb:gfp*$^{+}$ cells from *ddx41* mutants injected with control (left) or splice-blocking *cgas* morpholino (right) at 40 hpf. **P**. Graph depicting the absolute number of *nfκb:gfp*-high cells per embryo shown in O. Graphs display means ± stds with p-values calculated with unpaired Student's t-test (B,D,I,K,M) or a one-way ANOVA with Tukey's multiple testing correction (G,P), *=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001. N=2-3 replicates per experiment.

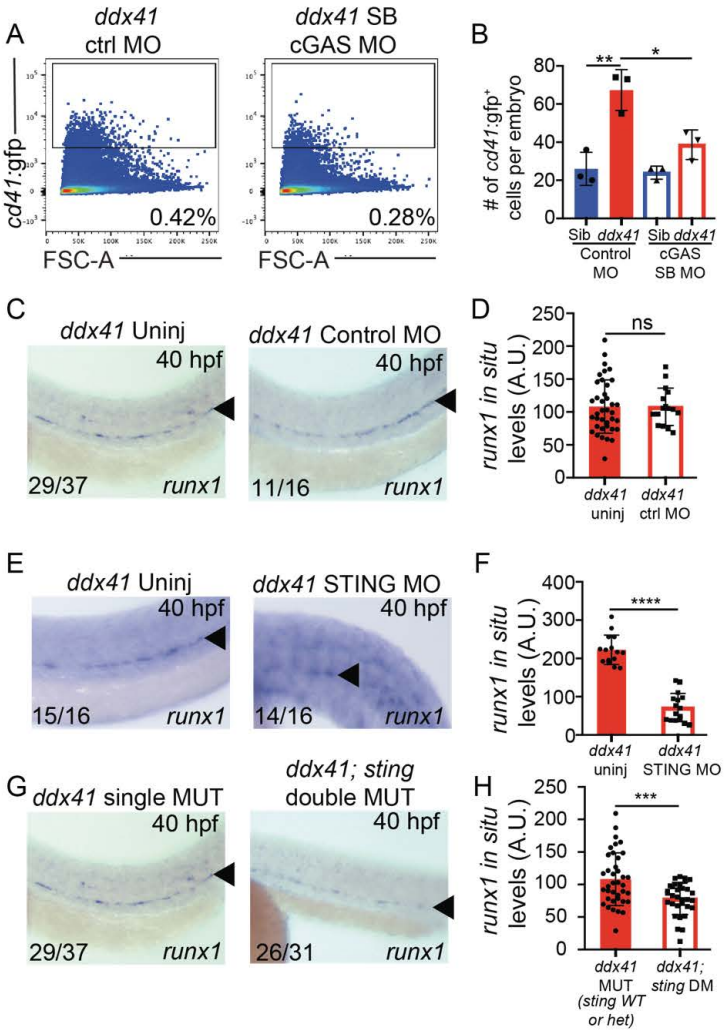


Figure S5. Ddx41 regulates HSPC number via the cGAS-STING inflammatory pathway. A. Flow cytometry plots of *cd41:*gfp^{+} HSPCs from *ddx41* mutants injected with control morpholino (left) and splice-blocking *cgas* morpholino (right) at 40 hpf. **B.** Graph depicting the absolute number of *cd41:*gfp^{+} HSPCs per embryo shown in A. **C,E,G.** *In situ* hybridization of the HSPC marker *runx1* at 40 hpf in respective control *ddx41* mutants (left) and experimental *ddx41* mutants (right). C: scrambled control morpholino, E: *sting* morpholino, and G: *sting*^{mk30} mutant. Numbers on bottom left corner indicate the fraction of embryos with the same phenotype as the one depicted in the image. A high magnification view of the AGM region is displayed. The aorta is marked with arrowheads. **D,F,H.** Quantification of *runx1 in situ* hybridization levels from C, E, and G, respectively.

Haematologica
HAEMATOL/2020/257246
Version 1

Ddx41 inhibition of DNA damage signaling permits erythroid progenitor expansion in zebrafish

Joshua T. Weinreb, Elianna S. Pechman, Rachel Weil, Varun Gupta, and Teresa V. Bowman

Disclosures: The authors declare no competing financial interests.

Contributions: JTW and TVB designed the project experimental approach. JTW, ESP, and RW performed the experiments. VG performed bioinformatics analysis. JTW and TVB analyzed the data. JTW and TVB wrote the manuscript; all authors reviewed and approved the manuscript.

Ddx41 inhibition of DNA damage signaling permits erythroid progenitor expansion in zebrafish

Joshua T. Weinreb[1,2], Elianna S. Pechman[1], Rachel Weil[1], Varun Gupta[3], Teresa V. Bowman[#1,2,4]

1. Albert Einstein College of Medicine, Department of Developmental and Molecular Biology, Bronx, NY, USA
2. Albert Einstein College of Medicine, Gottesman Institute for Stem Cell Biology and Regenerative Medicine, Bronx, NY, USA
3. Albert Einstein College of Medicine, Department of Cell Biology, Bronx, NY, USA
4. Albert Einstein College of Medicine, Department of Medicine (Oncology), Bronx, NY, USA

#To whom correspondence should be addressed:
Address: 1300 Morris Park Avenue, Bronx, NY 10461
Email: teresa.bowman@einsteinmed.org
Phone: (718) 430-4001
Fax: (718) 430-8567

Text word count: 3516

Abstract word count: 139

Number of figures: 6 main + 1 supplemental

Number of tables: 5

Number of references: 37

ABSTRACT

DEAD-box Helicase 41 (*DDX41*) is a recently identified factor mutated in hematologic malignancies whose function in hematopoiesis is unknown. Using an *in vivo* model of Ddx41 deficiency, we unveiled a critical role for this helicase in regulating erythropoiesis. We demonstrated that loss of *ddx41* leads to anemia caused by diminished proliferation of erythroid progenitors. Mis-expression and alternative splicing of cell cycle genes is rampant in *ddx41* mutant erythroid progenitors. We delineated that the DNA damage response is activated in mutant cells resulting in an Ataxia-telangiectasia mutated (ATM) and Ataxia-telangiectasia and Rad3-related (ATR)-triggered cell cycle arrest. Inhibition of these kinases partially suppressed *ddx41* mutant anemia. These findings establish a critical function for Ddx41 in promoting healthy erythropoiesis via protection from genomic stress and delineate a mechanistic framework to explore a role for ATM and ATR signaling in DDX41-mutant hematopoietic pathologies.

INTRODUCTION

Mutations in DEAD-box Helicase 41 (DDX41) were identified recently in hematologic malignancies including myelodysplastic syndrome (MDS), acute myeloid leukemia (AML), and acute erythroid leukemia (AEL)[1,2]. Germline *DDX41* frameshift mutations are loss-of-function and somatic missense mutations are thought to result in hypomorphic helicase activity[1,3]. The human genetics thus suggest that diminished function of this helicase is detrimental to hematopoiesis, but this has yet to be demonstrated in an animal model.

In particular, a significant number of *DDX41*-mutated MDS patients experience mild cytopenia in the years preceding diagnosis, indicating that anemia may be one of the first warning signs of disease[4]. Anemia in MDS is attributed to numerous cellular mechanisms including erythroid precursor apoptosis, defective progenitor expansion, and ineffective erythrocytic maturation[5-7]. The clinical findings suggest DDX41 could be important in erythropoiesis, but the cellular and molecular underpinnings remain unclear. Roles for DDX41 have been implicated in genomic stability, inflammation, and splicing, all processes linked to hematopoietic health, but the current lack of DDX41 mutant animal models have slowed exploration of its function in the blood system[1,8-11].

To uncover the *in vivo* role of DDX41 in erythropoiesis, we established a zebrafish *ddx41* loss-of-function mutant. We demonstrated that *ddx41* mutants develop anemia due to a decrease in erythroid progenitor expansion. Mechanistically, the anemia is due in part to ATM- and ATR-mediated cell cycle arrest induced by elevated DNA damage and mis-expression of cell cycle regulators. Our data demonstrate that Ddx41 plays a critical role in hematopoiesis and provide a possible mechanism by which anemia may arise in *DDX41*-mutated hematopoietic pathologies.

METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Zebrafish

Zebrafish were maintained as described[12]. All fish were maintained according to IACUC-approved protocols in accordance with Albert Einstein College of Medicine research guidelines. We utilized several established zebrafish lines. Genotyping and phenotyping to confirm their identity was performed for each new generation and for embryos. Mutants for *ddx41* (*ddx41*^{sa14887}) were acquired from the Zebrafish International Resource Center (ZIRC) as part of the Sanger Institute Zebrafish Mutation Project[13]. The mutation results in a premature stop codon at Tyrosine 410. As early as 24 hpf, there appears to be slight brain death and abnormal curvature of the tail in *ddx41* mutants, allowing for phenotypic identification of mutants compared to siblings based on morphology. Homozygous mutants are embryonic lethal, dying by 3dpf. This allows for the assessment of primitive and definitive hematopoietic development in zebrafish up to 3 dpf but precludes the assessment of hematopoiesis beyond this point of development. In addition to *ddx41* mutant, *gata1*:dsred[14] transgenics were used and scored for dsRED fluorescence within erythrocytes. The genotyping for *ddx41* was determined by mutant-specific PCR, which was initially verified by Sanger sequencing. Additionally, the genotype of *ddx41*$^{sa14887/+}$ carriers were validated by paired mating of heterozygous adult animals followed by examination of embryos for the expected homozygous recessive phenotype. Genomic DNA for genotyping was isolated by alkaline lysis[15]. The *ddx41*^{sa14887} mutant and wildtype-specific PCR products were analyzed by gel electrophoresis. Genotyping primers are provided here: *ddx41* wildtype Forward: GAGGCCAAGATGGTTTAT; *ddx41* wildtype Reverse: GCCTCCACACCTTTT AGCAG; *ddx41* MUT Forward: GCCTCCACACCTTTTAGCAG; *ddx41* MUT Reverse: GAGGCCAAGATGGTTTAA.

Drug treatments

All drugs were dissolved in DMSO. Dilutions were made in E3 embryo water. Concentrations used for experiments were 30nM for KU60019 (ATM inhibitor; Fisher Scientific) and 30nM for AZ20 (ATR inhibitor; Fisher Scientific). DMSO in E3 embryo water was used as the vehicle control.

Whole-mount *in situ* hybridization, o-dianisidine staining, and active caspase-3 staining

In situ hybridization steps were performed as described previously by Thisse *et al*[16] with minor modifications: before proteinase K permeabilization, embryos older than 28 hpf were bleached after re-hydration to remove pigmentation. The bleaching was done for 5-10 minutes using a bleaching solution of 0.8% KOH, 0.9% H_2O_2 and 0.1% Tween 20. Embryos were then scored manually, imaged and genotyped. The *β-globin*[17], *cmyb*[18], and *gata1*[19] probes were used, and *in situ* levels were quantified using FIJI in a similar manner to previously described methods[20].

O-dianisidine staining was performed as described previously[21]. Briefly, dechorionated live embryos were soaked in o-dianisidine staining solution (0.62 mg/mL o-dianisidine [Sigma-Aldrich], 10.9 μM sodium acetate, and 0.65% H_2O_2) for 10 minutes in the dark, then rinsed in phosphate-buffered solution plus 0.1% Tween 20 (PBT) twice and observed under the microscope. Afterward, embryos were preserved in a 4% paraformaldehyde (PFA) solution.

Flow cytometry

At 28 or 40 hpf, mutant and sibling embryos were binned based on morphological differences. For generation of single-cell suspensions, 10-20 embryos were first removed from their chorions using pronase (Roche), and then homogenized by manual dissociation using a sterile razor blade followed by digestion with Liberase (Roche). For the digestion, dissociated embryos were resuspended in 600μl 1× Dulbeccos-PBS (D-PBS) (Life Technologies) supplemented with a 1:65 dilution (9.23 μl) of 5 mg/ml Liberase and then incubated at 37°C for 7 minutes. The reaction was stopped with the addition of 5% (30μl) fetal bovine serum (FBS) (Life Technologies). The cells were then filtered through a 40-μm cell strainer (Falcon) and pelleted by centrifugation at 3000 rpm for 5 minutes. Cell pellets were resuspended in 400-700 μl FACS buffer (0.9× D-PBS, 5% FBS, 1% Penn/Strep (Life Technologies)). DAPI (4′,6-diamidino-2-phenylindole) was added to a final concentration of 1 μg/ml to facilitate exclusion of dead cells from the analysis. Samples were analyzed with a LSRII flow cytometer (BD Biosciences) and FlowJo software. Cells from non-fluorescent embryos were used to set gates above background. Quantification for the absolute number of cells was performed by acquiring all events in

a tube on the flow cytometer to determine the total number of target cells. This number was then divided by the total number of embryos analyzed to calculate the number of target cells per embryo.

Cell cycle and apoptosis analysis

For 5-ethynyl-2′-deoxyuridine (EdU) incorporation experiments, 28 hpf embryos were binned into 6 well plates, with ~30 embryos per well. Embryos were incubated with 20mM EdU for 2 hours. Generation of single-cell suspensions of 40 hpf embryos was accomplished as described above. EdU was detected using the Click-IT EdU Flow Cytometry Assay Kit according to the manufacturer's instructions (Invitrogen C10424). Flow cytometry analysis for active caspase-3 at 30 hpf was performed as previously described[22]. Samples were analyzed with a LSRII flow cytometer (BD Biosciences) and FlowJo software.

RNA-sequencing and splicing analysis

Erythrocytic progenitors from *ddx41* mutants and siblings at 40 hpf were isolated by FACS on a Mo-Flo cell sorter (Beckman Coulter). RNA from these cells was subsequently isolated using the Zymo Quick-RNA microprep kit (R2080) according to the manufacturer specification. The TURBO DNA-free kit (Life Technologies) was used for DNA removal following RNA extraction. RNA and library quality was assessed using a bioanalyzer Pico Chip (Agilent) by the Einstein Genomics core facility. Libraries for RNA samples with RIN≥9 were prepared by BGI using the Nugene low-input RNA sequencing kit. Five independent biological replicates were analyzed. On average, approximately 30 million paired-end 150 base pair sequencing reads were acquired per sample using the Illumina NovoSeq platform. DE-Seq2[23] was used to determine differential gene expression (defined as log2 fold change ≥ |1|, FDR $p<0.05$) between *ddx41* mutant and sibling control erythrocytes. Pathway analysis was analyzed using the Molecular Signatures Database (MSigDB)[24,25]. Differential inclusion rates of splicing events between *ddx41* mutants and siblings were calculated using rMATS[26] (replicate Multivariate Analysis of Transcript Splicing). Splicing events with inclusion differences ≥ 10% and FDR ≤ 0.01 were

124 considered significantly altered. Pathway analysis for alternatively spliced genes was also
125 performed using MSigDB.
126
127 **Single-cell immunofluorescence of zebrafish embryonic cells**
128 Single-cell suspensions were prepared as described above, and cell staining was
129 performed as described in Sorrells & Nik *et al.*[22]. The rabbit polyclonal anti-zebrafish
130 γH2AX primary antibody (GeneTex GTX127342) was used at a dilution of 1:500. All
131 antibodies were diluted in blocking buffer (5% bovine serum albumin (BSA)/0.2%
132 milk/PBS) and incubated for 3 h at room temperature followed by 3 washes in 0.1%
133 Tween 20/PBS. The goat anti-rabbit AlexaFluor-594 (Thermo Fisher Scientific) was used
134 at a dilution 1:1000 and incubated for 1 h at room temperature followed by 3 washes in
135 0.1% Tween 20/PBS. After final washes, cells were mounted with DAPI-Fluoromount-G
136 (Southern Biotech), covered with a 25×25 mm glass coverslip and sealed with clear nail
137 polish. Fluorescence intensity measurements of γH2AX were performed using FIJI.
138
139 **May-Grunwald Giemsa staining of primitive erythroid cells**
140 This protocol is adopted as described by[27]. Two-day old embryos were placed on poly-L-
141 lysine coated slides in a drop of 1× D-PBS + 1% bovine serum albumin (Sigma). Blood
142 cells were released from the embryo by puncturing the pericardial sac and upper yolk sac
143 with fine forceps. The slides were air-dried at room temperature prior to staining. For
144 staining, slides were immersed in undiluted May-Grunwald stain (Eng Scientific May-
145 Grunwald stain solution 1, Fisher Scientific) for 2 minutes and briefly rinsed in ddH_20.
146 Slides were then immersed in diluted Giemsa stain (diluted 1:4 with milliQ water) for 20
147 minutes (Eng Scientific May-Grunwald stain solution II, Fisher Scientific) and briefly rinsed
148 in ddH_20. Once slides were dry, a drop of Permount solution (Fisher Scientific) was added
149 and slides were covered with cover slips and left overnight to dry. Once slides were dry,
150 the cells were visualized with a 63× oil-immersion lens.
151
152 **Image acquisition and processing**
153 Brightfield and fluorescent images were acquired using a Zeiss Axio Discovery.V8 with
154 an AxioCam HRc camera. Confocal images for zebrafish single-cell immunofluorescence

were acquired using a Leica SP5 AOBS Inverted DMI6000 microscope with a 63X oil objective and zoomed in fields were taken at 6.5X zoom. FIJI/ImageJ was used for quantification of images[28]. Flowjo, GraphPad Prism 8, Adobe Photoshop and Adobe Illustrator were used to generate figures.

Statistics

Experiments were performed with a minimum of three replicates. Statistical analyses were performed as indicated in each figure using unpaired Student's t-test or a one-way ANOVA with Tukey's multiple testing correction as appropriate; error bars indicate the standard deviation of mean, unless otherwise indicated.

RESULTS

Ddx41 regulates erythroid development

To explore a function for Ddx41 in hematopoiesis, we examined erythrocyte formation and differentiation in zebrafish *ddx41* loss-of-function mutants (*ddx41*^{sa14887}). DDX41 is highly conserved between humans and zebrafish suggesting that lessons learned about the *in vivo* role of zebrafish *ddx41* function in hematopoiesis will be relevant to human *DDX41*. In zebrafish, primitive hematopoiesis begins ~12-24 hours post fertilization (hpf), producing embryonic erythrocytes and myeloid cells that constitute the hematopoietic system early on in development[29,30]. Cells of the erythrocytic lineage first arise from the inner cell mass (ICM) within the posterior lateral mesoderm (PLM)[19]. These erythrocytes express factors such as the progenitor transcription factor *c-myb* and the erythroid-specific transcription factor *gata1* required for erythropoiesis[31]. Using *in situ* hybridization, we determined that these erythroid progenitor markers were expressed similarly in *ddx41* mutants compared to siblings at 22hpf, indicating initial erythroid specification is unaffected (Figures 1A-D). Oxygenated hemoglobinized erythrocytes are detectable beginning around 36 hpf using o-dianisidine staining[32]. In *ddx41* mutants, we observed little o-dianisidine-positive erythroid cells at 36 hpf (Figure 1E). Mutants display some developmental delay that becomes more severe as the embryos get older, thus we also

185 looked at o-dianisidine staining at 48 hpf and noted little to no improvement to oxygenated
186 hemoglobin levels (Figures 1F-G).
187
188 This decrease in oxygenated hemoglobin levels can be driven by a defect in maturation
189 or proliferation. To assess maturation, we analyzed the morphology of isolated
190 erythrocytes using May-Grunwald-Giemsa staining. While the *ddx41* mutant erythrocytes
191 did display a mild megaloblastoid-like phenotype, the majority of erythrocytes in both
192 *ddx41* mutants and siblings were at the orthochromatophilic erythroblast stage,
193 suggesting there was not a major block to differentiation in the mutants (Figure 1H).
194
195 To acquire enough erythrocytes to perform the morphological assessment, we needed to
196 bleed four times as many *ddx41* mutant embryos as compared to sibling control embryos,
197 suggesting mutants had fewer erythrocytes than siblings. To test this hypothesis, we
198 assessed the number of *gata1*:dsred^{+} erythroid progenitors in *ddx41* mutants and siblings
199 using flow cytometry quantification. We determined that the absolute number of
200 *gata1*:dsred^{+} erythrocytes per embryo was significantly reduced in *ddx41* mutants
201 compared to siblings at both 28 and 40 hpf (Figures 2A-D). Taken together, these data
202 indicate that *ddx41* mutants develop anemia due to a decrease in erythroid progenitor cell
203 number.
204
205 Erythroid progenitors arising from both primitive and definitive erythroid-myeloid
206 progenitor (EMP)-derived waves are present during the developmental time points
207 analyzed. To determine whether there were defects in EMP-derived erythropoiesis, we
208 performed *in situ* hybridization for the progenitor marker *c-myb* at 26 and 36 hpf
209 and *gata1* at 26 hpf in siblings and *ddx41* mutants (Figures 2E-F and S1A-D). Expression
210 of both *c-myb* and *gata1* within the posterior blood island (PBI) region where EMPs form
211 were not decreased and in fact were increased in *ddx41* mutants as compared to siblings.
212 These data indicate that the reduction in erythrocytes observed up to 48 hpf
213 in *ddx41* mutants are mostly due to defects in primitive wave-derived erythroid
214 progenitors.
215

Cell cycle genes are mis-expressed and alternatively spliced in *ddx41* mutant erythroid progenitors

To mechanistically assess the underlying cause of the erythrocytic defect in *ddx41* mutants, we conducted RNA-sequencing (RNA-seq) on *gata1:*dsred^{+} erythrocytes isolated from *ddx41* mutants and siblings at 40 hpf. Over 1800 genes were downregulated and more than 1900 were upregulated in *ddx41* mutants compared to siblings (Figure 3A, Table S1, log2 fold change $\geq |1|$, adj. p-value < 0.05). To understand if particular pathways were enriched in the differentially expressed genes, we performed gene set level analysis on the upregulated and downregulated gene lists by comparing each to the Molecular Signature Database (MSigDb), a platform that computes overlaps between classes of genes that are over- or under-represented in a lists of genes in known pathways[24,25]. In the downregulated gene list, mRNA splicing was the top gene set with DNA replication, cell cycle, and DNA repair also enriched (Figure 3B, Table S2). In the upregulated gene list, genes associated with adaptive immunity, post-translational modifications, innate immune system, and cell cycle were enriched (Figure 3C, Table S3).

Ddx41 interacts with components of the spliceosome[1]. Additionally, the top downregulated pathway in our gene set was pre-mRNA splicing, thus we examined how *ddx41* loss affected mRNA splicing in erythrocytes. When comparing splicing between *ddx41* mutants and siblings, a total of 370 alternative splicing events were observed (Figures 3D-E, Table S4). The specific splicing defects detected included exon skipping, which was the most frequently altered splicing event, intron retention, alternative 5'-splice site usage, alternative 3'-splice site usage, and changes in mutually exclusive exon usage. The changes were found almost equally in *ddx41* mutants and siblings suggesting that Ddx41 functions in a diverse range of splicing events, potentially as both a splicing enhancer and suppressor. This pattern is in contrast to mutants in core spliceosomal components such as *sf3b1* that show highly mutant-centric alternative splicing patterns[33]. Finally, we compared the list of alternatively spliced genes in *ddx41* mutants and siblings to the MSigDB to assess if particular pathways were affected and determined that cell cycle, mRNA splicing, innate immunity, and TP53 regulation of DNA repair genes were

among the top enriched pathways (Figure 3F, Table S5). These results depict that Ddx41 influences the expression and splicing of cell cycle and DNA repair genes in erythrocytes.

Ddx41 deficiency triggers cell cycle arrest in erythroid progenitors

The diminished number of erythroid progenitors and dysregulated expression of cell cycle genes suggests that defects in erythrocyte proliferation could contribute to the anemia in *ddx41* mutants. To examine proliferation, we analyzed cell cycle status of 30 hpf *gata1:*dsred^{+} erythroid progenitors by flow cytometry quantification of DNA synthesis via 5-Ethynyl-2'-deoxyuridine (EdU) incorporation and DNA content via DAPI incorporation. The *ddx41* mutant *gata1:*dsred^{+} progenitors showed a reduction of cells in S phase and an accumulation of cells in the G0/G1 and G2/M phases compared to sibling controls (Figures 4A-B). These results are in-line with a decrease in proliferation in *ddx41* mutant erythrocytes caused by cell cycle arrests at the G0/G1-to-S phase and G2-to-M transitions. Prolonged cell cycle arrests can lead to apoptosis[34], thus we also assessed apoptosis in *ddx41* mutants. We measured levels of active caspase-3, an essential executor of apoptosis, in *ddx41* mutant and sibling *gata1:*dsred^{+} erythrocytes by flow cytometry. We observed a significant increase in active caspase-3 in *ddx41* mutant *gata1:*dsred^{+} erythrocytes at 30 hpf (Figures 4C-D). These data indicate that both cell cycle arrest and elevated apoptosis in *ddx41* mutant erythrocytes may drive anemia.

Ddx41 regulation of ATM and ATR signaling contributes to proper erythropoiesis

These cellular phenotypes coupled with the decreased expression of DNA repair genes in *ddx41* mutants imply that loss of *ddx41* could promote DNA damage. To address this question, we analyzed the DNA damage response (DDR) by performing immunofluorescence (IF) for γH2AX in *ddx41* mutants and siblings. We showed that γH2AX levels were increased nearly two-fold in *ddx41* mutants compared to sibling cells (Figures 4E-F). These data demonstrate that Ddx41 regulates DDR *in vivo*.

Our model is that loss of Ddx41 contributes to excessive DDR signaling and subsequent cell cycle arrest in erythrocytes, leading to anemia in *ddx41* mutants. If correct, then inhibiting components of the DDR pathway would 1) reverse cell cycle defects and 2)

increase erythrocyte levels. To test this model, we examined how the two primary mediators of DDR, Ataxia-telengiectasia-mutated (ATM) and Ataxia-telengiectasia and Rad3-related (ATR), affected erythrocytic cell cycle kinetics in *ddx41* mutants. We assessed cell cycle status of 30 hpf *ddx41*-mutant *gata1*:dsred^{+} erythroid progenitors in embryos treated with DMSO vehicle control, the ATM inhibitor KU60019, or the ATR inhibitor AZ20. There was a significant increase of *gata1*:dsred^{+} *ddx41*-mutant cells in S phase when treated with either ATM or ATR inhibitors as compared to DMSO vehicle control (Figures 5A-B). Additionally, pharmacological inhibition of ATM or ATR both increased erythropoietic output in *ddx41* mutants, as measured by quantification of *gata1*:dsred^{+} erythrocyte numbers per embryo using flow cytometry (Figures 5C-D). Taken together, these data indicate that DDR signaling triggers a G0/G1 cell cycle arrest in *ddx41*-mutant erythrocytes that results in a reduction of erythroid progenitor cell number.

Finally, we wanted to assess if increasing the number of erythroid progenitors via ATM or ATR inhibition would increase the number of oxygenated erythrocytes in *ddx41* mutants. Surprisingly, we only observed a significant increase in o-dianisidine-positive erythrocytes in *ddx41* mutants treated with ATM inhibitor, but not ATR inhibitor (Figures 6A-B). These data indicate that Ddx41 regulation of ATM might have a broader impact on erythropoiesis than ATR signaling.

DISCUSSION

Although *DDX41* mutations are found in numerous human hematologic disease, its function in hematopoiesis is unknown. Our work is the first to establish Ddx41 as a critical mediator of erythropoiesis with *ddx41* loss showing a profound effect on the expression of cell cycle and DNA repair genes in erythrocytes. Subsequently, we demonstrated that *ddx41* mutant erythroid progenitors undergo cell cycle arrest. The DNA damage response is elevated in *ddx41* cells and triggers an ATM and ATR-triggered cell cycle arrest. Inhibition of ATM and ATR partially suppressed anemia in *ddx41* mutants. These findings establish Ddx41 as a positive regulator of erythropoiesis by preventing genomic stress and promoting proper erythroid progenitor expansion.

DDX41 was previously identified as a mediator of genomic stability in a cell line-based genome-wide siRNA screen[11]. However, a role for DDX41 in genomic integrity as well as the downstream consequences of its loss were never demonstrated *in vivo*. Our current work reveals that Ddx41 regulates genomic integrity *in vivo*, and that loss of *ddx41* leads to both cell cycle arrest and apoptosis in erythrocytes that contributes to anemia in *ddx41* mutants. We establish that ATM and ATR signaling contribute to these attributes, but only ATM inhibition significantly increased the number of o-dianisidine-positive erythrocytes in *ddx41* mutants. The differential impact on oxygenated erythrocyte output by inhibition of ATM and ATR might indicate that Ddx41-regulated ATM signaling is more critical for proper erythropoiesis. However, it should be noted that though the ATM and ATR kinases respond uniquely, their exists an extensive 'cross-talk' between them which can make determining which precise pathway is involved in a phenotype confusing[35]. Further dissection of the role of DDX41 in ATM and ATR pathway regulation will need to be investigated.

Splicing mutations are commonly found in hematologic malignancies[36,37]. DDX41 interacts with multiple components of the spliceosome[1]. Our work aligns with prior studies showing *DDX41* insufficiency associates with numerous deleterious splicing outcomes. If and how these splicing events contribute to hematopoietic pathogenesis is unclear. We show that components related to cell cycle and DNA repair are commonly mis-spliced in *ddx41* mutants. Therefore, it is possible that loss of *ddx41* may be mediating cell cycle arrest and activation of DDR via mis-splicing of crucial regulators of these pathways. The contribution of DDR pathway component mis-splicing in human cytopenias remains to be addressed.

In sum, our study unveils a critical role for Ddx41 as a key gatekeeper to maintain cell cycle progression, a necessary component for erythrocytic development. We demonstrated that deficiency of *ddx41* triggers cell cycle arrest via activation of ATM and ATR, which ultimately mediates a decrease in proliferation and maturation of erythrocytic progenitors in *ddx41* mutants. These findings establish a critical function for Ddx41 in

promoting healthy erythropoiesis by suppressing genomic stress and present a potential role for ATM and ATR signaling in DDX41-mutant pathologies.

ACKNOWLEDGMENTS

This work was funded by American Cancer Society RSG-129527-DDC, DOD BM180109, 1R56DK121738-01 and the Evans MDS Foundation (to TVB), MSTP training grant T32GM007288-45 and F30 fellowship 1F30HL142161 (to JTW), and 1R01GM057829-23 to Charles Query for support of VG.

We also want to acknowledge the assistance of numerous core facilities at Albert Einstein College of Medicine including Flow Cytometry, Analytical Imaging, and Genomics Facilities (funded by NCI Cancer Grant P30CA013330), and the Zebrafish Core Facility.

AUTHORSHIP CONTRIBUTIONS

JTW and TVB designed the project experimental approach. JTW, ES, and RW performed the experiments. VG performed bioinformatics analysis. JTW and TVB analyzed the data. JTW and TVB wrote the manuscript; all authors reviewed and approved the manuscript.

CONFLICT OF INTEREST DISCLOSURES

The authors declare no competing financial interests.

359 **FIGURE LEGENDS**

360 **Figure 1. Loss of *ddx41* causes anemia. A,C.** *In situ* hybridization of the erythroid
361 markers *cmyb* (**A**) and *gata1* (**C**) at 22 hours post fertilization (hpf) in sibling controls (left)
362 and *ddx41* mutants (right). Arrowheads highlight the inner cell mass (ICM) region in the
363 embryos. **B,D.** Quantification of *c-myb* (**B**) and *gata1* (**D**) *in situ* hybridization levels from
364 (A) and (C), respectively. Quantification was done using Fiji. **E,F**. Staining for o-
365 dianisidine, marking functional hemoglobin in mature primitive erythrocytes, in sibling
366 controls (left) and *ddx41* mutants (right) at 36 hpf (E) and 48hpf (F). Numbers on bottom
367 left corner indicate the fraction of embryos with the same phenotype as the one depicted
368 in the image. **G.** Graph depicting frequency of designated o-dianisidine staining levels in
369 sibling controls and *ddx41* mutants at 48 hpf. **H**. Representative images of
370 orthochromatophilic erythroblasts stained with May–Grunwald–Giemsa from sibling
371 controls (left) and *ddx41* mutants (right) at 48hpf. Scale bar=20um. Graphs display means
372 ± standard deviations (stds) with p-values calculated with unpaired Student's t-test,
373 ns=not significant ($p>0.05$), ****=$p<0.0001$. For *in situs* and o-dianisidine staining n=6-72
374 embryos per experiment.

375

376 **Figure 2. Ddx41 regulates erythroid progenitor numbers. A,C.** Flow cytometry plots
377 of *gata1*:dsred$^+$ erythroid cells from sibling controls (left) and *ddx41* mutants (right) at 28
378 hpf (A) and 40 hpf (C). **B,D**. Graphs depicting the absolute number of *gata1*:dsred$^+$
379 erythroid cells per embryo from (A) and (C), respectively. **E.** *In situ* hybridization of *cmyb*
380 at 26 hpf in sibling controls (left) and *ddx41* mutants (right). **F.** Quantification of *cmyb in*
381 *situ* hybridization levels from (E). Quantification was done using Fiji. Graphs display
382 means ± standard deviations (stds) with p-values calculated with unpaired Student's t-
383 test, *=$p<0.05$, **=$p<0.01$, ***=$p<0.001$. For flow cytometry, n=5 pools of ~5-20 embryos
384 per pool; for *in situ* hybridization, n=10-12 embryos.

385

386 **Figure 3. Cell cycle genes are mis-expressed and alternatively spliced in *ddx41***
387 **mutant erythroid progenitors. A**. Volcano plot displaying differentially expressed genes
388 between *gata1*:dsred$^+$ erythrocytes from *ddx41* mutants and siblings. Significant
389 differences are defined as FDR <0.05 and log2 fold change ≥ 1. Black vertical lines denote

the fold-change threshold and the black horizontal line denotes the FDR threshold. Five biological replicates for both *ddx41* mutants and siblings were used to generate RNA-sequencing data. **B-C**. Representative charts of pathways significantly enriched in genes downregulated (B) or upregulated (C) in *ddx41* mutant erythroid progenitors compared to sibling controls as determined by MSigDB analysis. **D.** Schematic of the different splicing events detected by analysis with rMATS. Significant differences are defined as FDR ≤0.01 and $\Delta \psi \geq |0.1|$. The percentage of alternatively spliced events in mutants is based on the total number of altered splicing events in *ddx41* mutants as compared to siblings. Five replicates for both *ddx41* mutants and siblings were used to generate RNA-sequencing data. **E.** Graph depicting the $\Delta \psi$ of individual splicing events between sibling controls and *ddx41* mutants. **F.** Representative charts of pathways significantly enriched in alternatively spliced genes in *ddx41* mutant erythrocytes compared to sibling controls as determined by MSigDB analysis. SE- skipped exons, RI- retained introns, A5SS- alternative 5' splice site, A3SS- alternative 3' splice site, and MXE- mutually exclusive exons.

Figure 4. Ddx41 deficiency triggers cell cycle arrest and DNA damage response in erythroid progenitors. A. Cell cycle analysis of *gata1*:dsred^{+} erythroid cells from sibling controls (left) and *ddx41* mutants (right) after a 2-hour pulse of EdU at 28 hpf. EdU incorporation (y-axis) and DAPI content (x-axis) were measured by flow cytometry at 30 hpf. **B.** Quantification of the percentage of cells in each cell cycle phase from (A). **C.** Flow cytometry analysis of active-caspase 3 in *gata1*:dsred^{+} erythroid cells from sibling controls (left) and *ddx41* mutants (right). **D.** Quantification of the percentage of *gata1*:dsred^{+} erythroid cells that are active caspase-3-positive (C). **E.** Confocal images showing immunofluorescence of nuclei (DAPI) and γH2AX in cells isolated from 28 hpf siblings (top) and *ddx41* mutants (bottom). Scale bar=5μm. **F.** Quantification of γH2AX levels from (E). Graphs display means ± standard deviations (stds) with p-values calculated with unpaired Student's t-test, *=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001. For flow cytometry, n=3-5 pools of ~5-20 embryos per pool.

Figure 5. Ddx41 regulation of ATM/ATR signaling contributes to proper erythroid progenitor proliferation. A. Cell cycle analysis of *gata1*:dsred^{+} erythroid cells from *ddx41* mutants treated with DMSO (left), 30nM KU60019 (ATM inhibitor, middle), and 30nM AZ20 (ATR inhibitor, right) after a 2-hour pulse of EdU at 28 hpf. EdU incorporation (y-axis) and DAPI content (x-axis) were measured by flow cytometry at 30 hpf. **B.** Quantification of the percentage of cells in each cell cycle phase from (A). **C.** Flow cytometry plots of *gata1*:dsred^{+} erythroid cells from *ddx41* mutants treated with DMSO (left), 30nM KU60019 (ATM inhibitor, middle), and 30nM AZ20 (ATR inhibitor, right). **D**. Graphs depicting the absolute number of *gata1*:dsred^{+} erythroid cells per embryo from (C). Graphs display means ± standard deviations (stds) with p-values calculated with a one-way ANOVA with Tukey's multiple testing correction, *=p<0.05, **=p<0.01, ***=p<0.001. For flow cytometry, n=3-5 pools of ~5-20 embryos per pool.

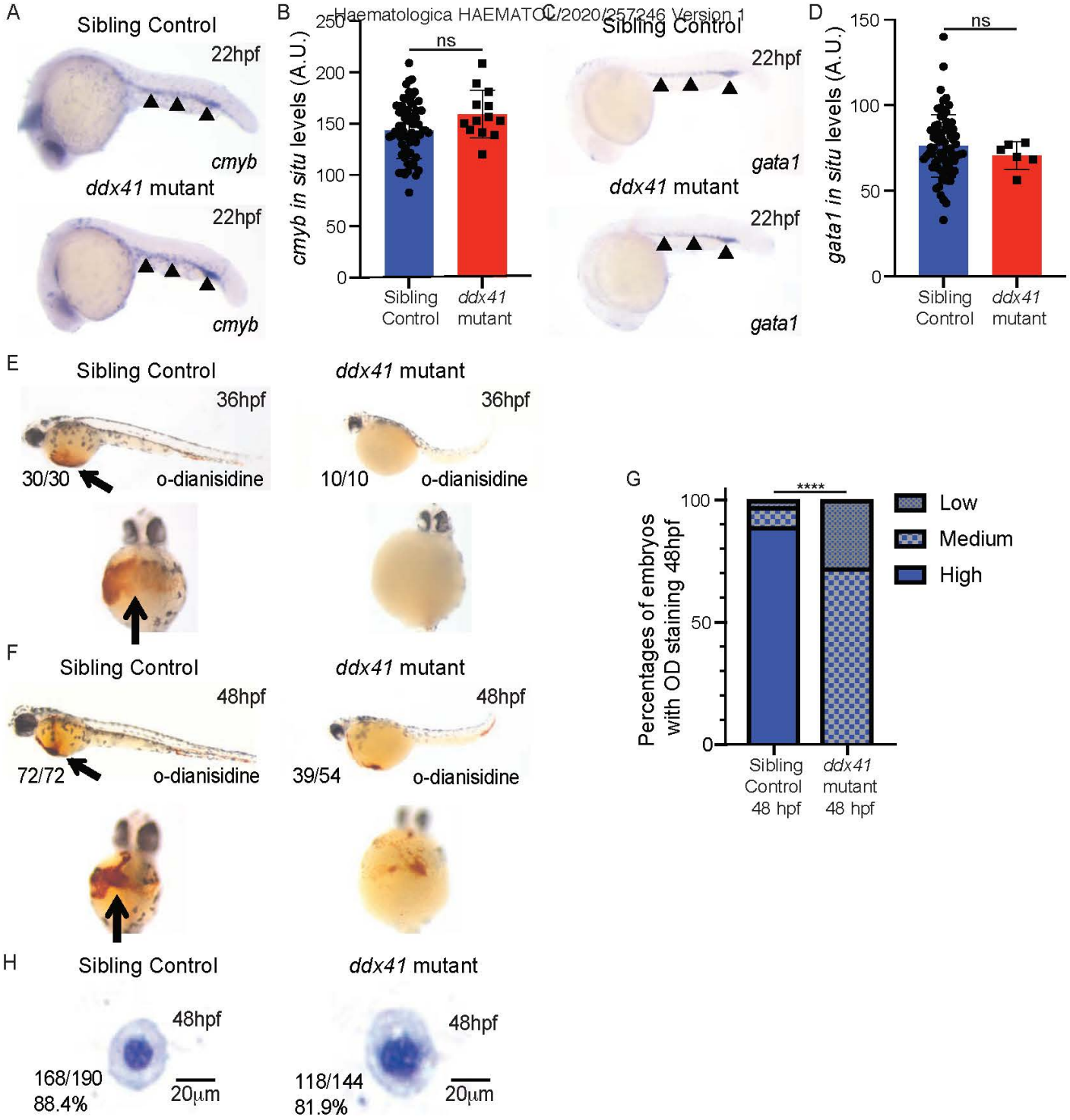
Figure 6. ATM inhibition partially suppresses *ddx41* mutant anemia. A. Representative images of o-dianisidine staining and corresponding levels of staining from *ddx41* `mutants treated with DMSO (top), 30nM KU60019 (ATMi, middle), and 30nM AZ20 (ATRi, bottom). Numbers on bottom left corner indicate the fraction of embryos with the same phenotype as the one depicted in the image. **B.** Graph depicting frequency of designated o-dianisidine staining levels in sibling controls and *ddx41* mutants at 40 hpf treated with DMSO vehicle control, 30nM KU60019 (ATMi), and 30nM AZ20 (ATRi). Graphs display means ± standard deviations (stds) with p-values calculated with a one-way ANOVA with Tukey's multiple testing correction, *=p<0.05, ****=p<0.0001. For o-dianisidine staining, n=36-67 embryos per experiment.

REFERENCES

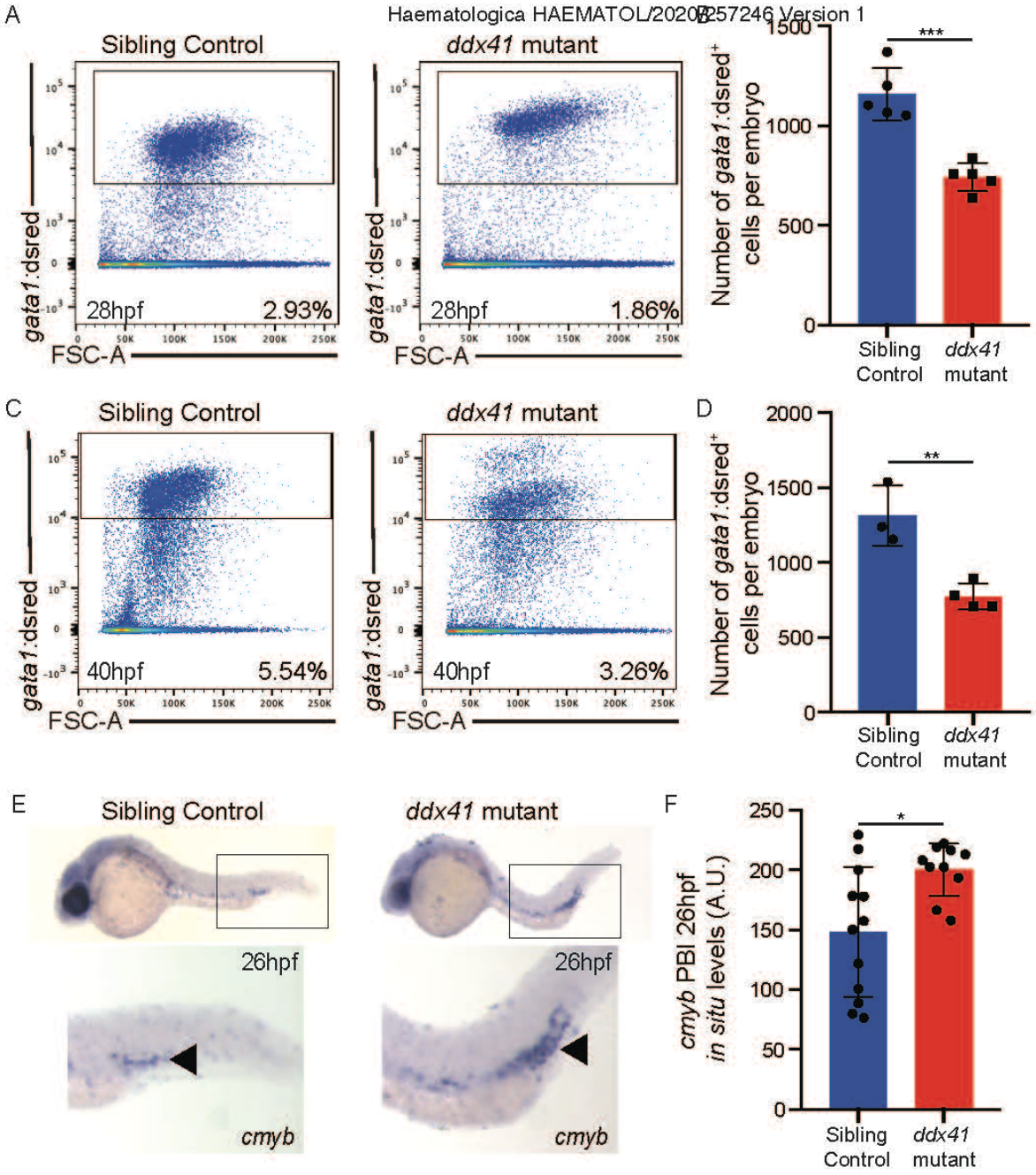
1. Polprasert C, Schulze I, Sekeres MA, et al. Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms. *Cancer Cell*. 2015;27(5):658-670.
2. Iacobucci I, Wen J, Meggendorfer M, et al. Genomic subtyping and therapeutic targeting of acute erythroleukemia. *Nat Genet*. 2019;51(4):694-704.
3. Yoneyama-Hirozane M, Kondo M, Matsumoto SI, et al. High-Throughput Screening to Identify Inhibitors of DEAD Box Helicase DDX41. *SLAS Discov*. 2017:2472555217705952.
4. Sebert M, Passet M, Raimbault A, et al. Germline DDX41 mutations define a significant entity within adult MDS/AML patients. *Blood*. 2019.
5. Schneider RK, Schenone M, Ferreira MV, et al. Rps14 haploinsufficiency causes a block in erythroid differentiation mediated by S100A8 and S100A9. *Nat Med*. 2016;22(3):288-297.
6. Danilova N, Sakamoto KM, Lin S. Ribosomal protein L11 mutation in zebrafish leads to haematopoietic and metabolic defects. *Br J Haematol*. 2011;152(2):217-228.
7. Payne EM, Virgilio M, Narla A, et al. L-Leucine improves the anemia and developmental defects associated with Diamond-Blackfan anemia and del(5q) MDS by activating the mTOR pathway. *Blood*. 2012;120(11):2214-2224.
8. Parvatiyar K, Zhang Z, Teles RM, et al. The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. *Nat Immunol*. 2012;13(12):1155-1161.
9. Zhang Z, Yuan B, Bao M, Lu N, Kim T, Liu YJ. The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nat Immunol*. 2011;12(10):959-965.
10. Zhang Z, Bao M, Lu N, Weng L, Yuan B, Liu YJ. The E3 ubiquitin ligase TRIM21 negatively regulates the innate immune response to intracellular double-stranded DNA. *Nat Immunol*. 2013;14(2):172-178.
11. Paulsen RD, Soni DV, Wollman R, et al. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell*. 2009;35(2):228-239.
12. Lawrence C. Advances in zebrafish husbandry and management. *Methods Cell Biol*. 2011;104:429-451.
13. Kettleborough RN, Busch-Nentwich EM, Harvey SA, et al. A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature*. 2013;496(7446):494-497.
14. Traver D, Paw BH, Poss KD, Penberthy WT, Lin S, Zon LI. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol*. 2003;4(12):1238-1246.

15. Meeker ND, Hutchinson SA, Ho L, Trede NS. Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *Biotechniques*. 2007;43(5):610, 612, 614.
16. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc*. 2008;3(1):59-69.
17. Brownlie A, Hersey C, Oates AC, et al. Characterization of embryonic globin genes of the zebrafish. *Dev Biol*. 2003;255(1):48-61.
18. Liao EC, Paw BH, Oates AC, Pratt SJ, Postlethwait JH, Zon LI. SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish. *Genes Dev*. 1998;12(5):621-626.
19. Detrich HW, 3rd, Kieran MW, Chan FY, et al. Intraembryonic hematopoietic cell migration during vertebrate development. *Proc Natl Acad Sci U S A*. 1995;92(23):10713-10717.
20. Dobrzycki T, Krecsmarik M, Bonkhofer F, Patient R, Monteiro R. An optimised pipeline for parallel image-based quantification of gene expression and genotyping after in situ hybridisation. *Biol Open*. 2018;7(4).
21. Lieschke GJ, Oates AC, Crowhurst MO, Ward AC, Layton JE. Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. *Blood*. 2001;98(10):3087-3096.
22. Sorrells S, Nik S, Casey M, et al. Spliceosomal components protect embryonic neurons from R-loop-mediated DNA damage and apoptosis. *Dis Model Mech*. 2018.
23. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
24. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
25. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425.
26. Shen S, Park JW, Lu ZX, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A*. 2014;111(51):E5593-5601.
27. De La Garza A, Cameron RC, Nik S, Payne SG, Bowman TV. Spliceosomal component Sf3b1 is essential for hematopoietic differentiation in zebrafish. *Exp Hematol*. 2016;44(9):826-837 e824.
28. Collins TJ. ImageJ for microscopy. *Biotechniques*. 2007;43(1 Suppl):25-30.

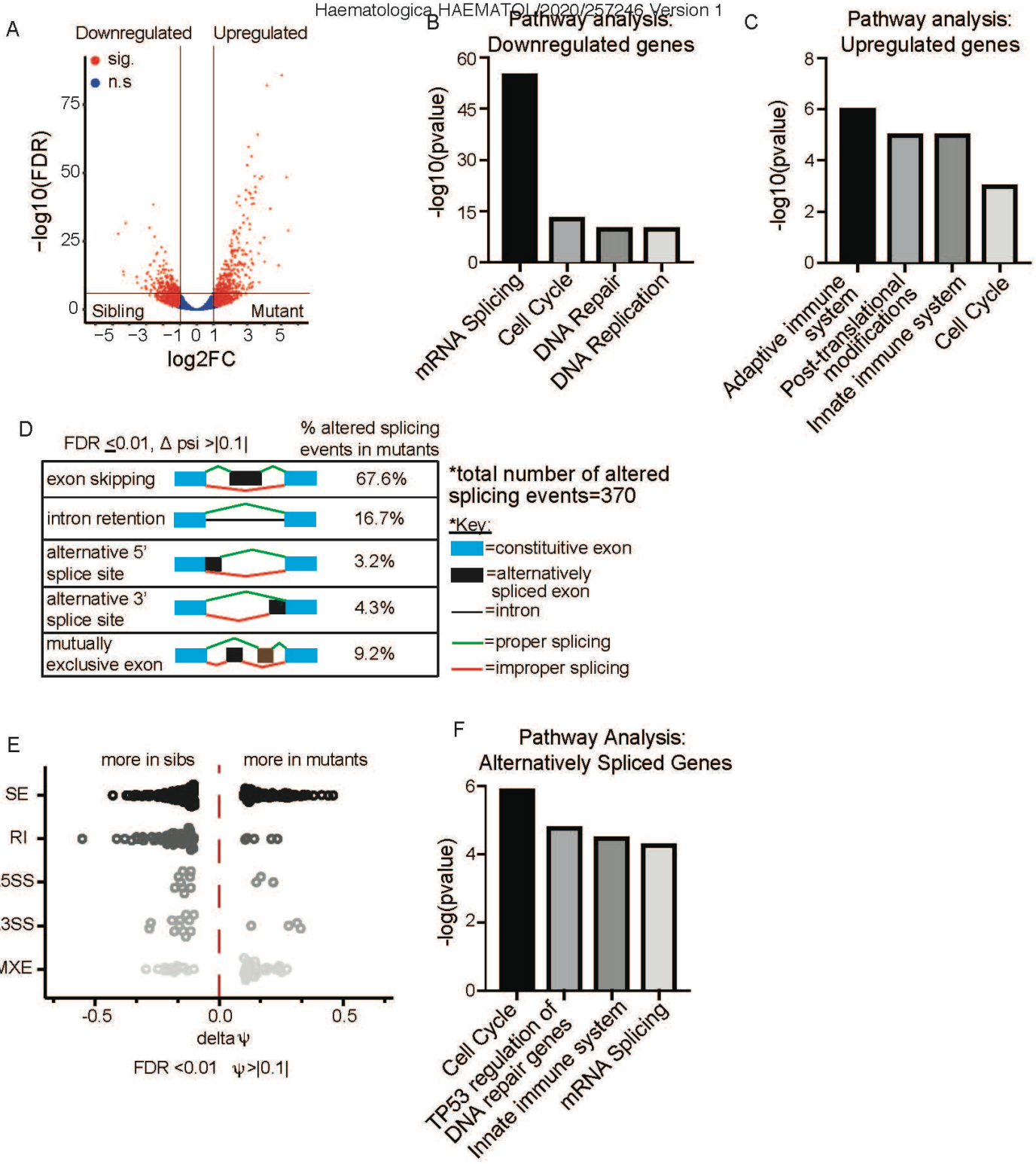
510 29. Carroll KJ, North TE. Oceans of opportunity: exploring vertebrate hematopoiesis in
511 zebrafish. *Exp Hematol*. 2014;42(8):684-696.
512 30. Clements WK, Traver D. Signalling pathways that control vertebrate haematopoietic stem
513 cell specification. *Nat Rev Immunol*. 2013;13(5):336-348.
514 31. Lyons SE, Lawson ND, Lei L, Bennett PE, Weinstein BM, Liu PP. A nonsense mutation in
515 zebrafish gata1 causes the bloodless phenotype in vlad tepes. *Proc Natl Acad Sci U S A*.
516 2002;99(8):5454-5459.
517 32. Paffett-Lugassy NN, Zon LI. Analysis of hematopoietic development in the zebrafish.
518 *Methods Mol Med*. 2005;105:171-198.
519 33. De La Garza A, Cameron RC, Gupta V, Fraint E, Nik S, Bowman TV. The splicing factor
520 Sf3b1 regulates erythroid maturation and proliferation via TGFbeta signaling in zebrafish. *Blood*
521 *Adv*. 2019;3(14):2093-2104.
522 34. Orth JD, Loewer A, Lahav G, Mitchison TJ. Prolonged mitotic arrest triggers partial
523 activation of apoptosis, resulting in DNA damage and p53 induction. *Mol Biol Cell*.
524 2012;23(4):567-576.
525 35. Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell*
526 *Biol*. 2008;9(8):616-627.
527 36. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing
528 machinery in myelodysplasia. *Nature*. 2011;478(7367):64-69.
529 37. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients
530 with myelodysplastic syndromes. *Leukemia*. 2014;28(2):241-247.
531



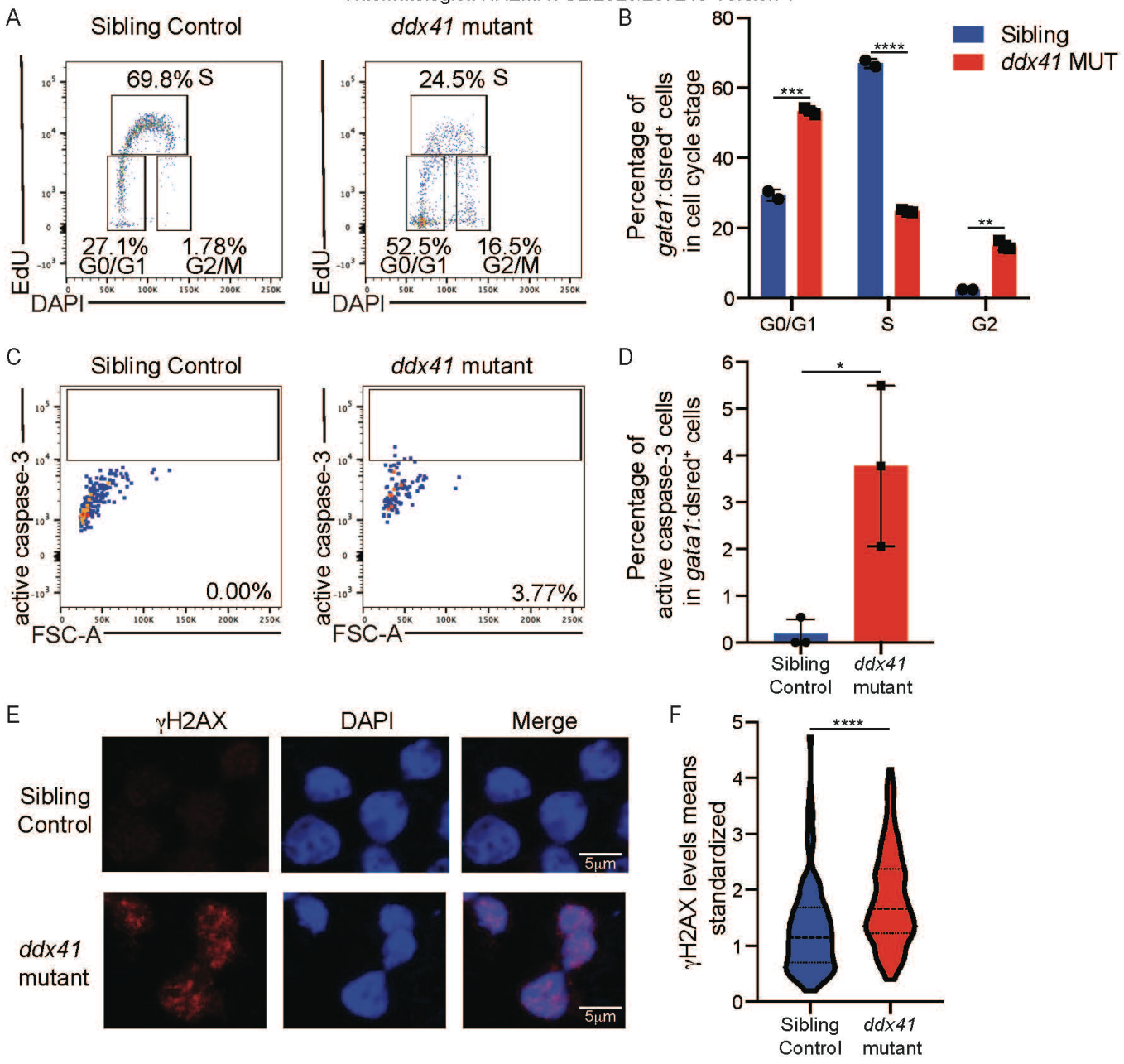
Weinreb et al. Figure 1



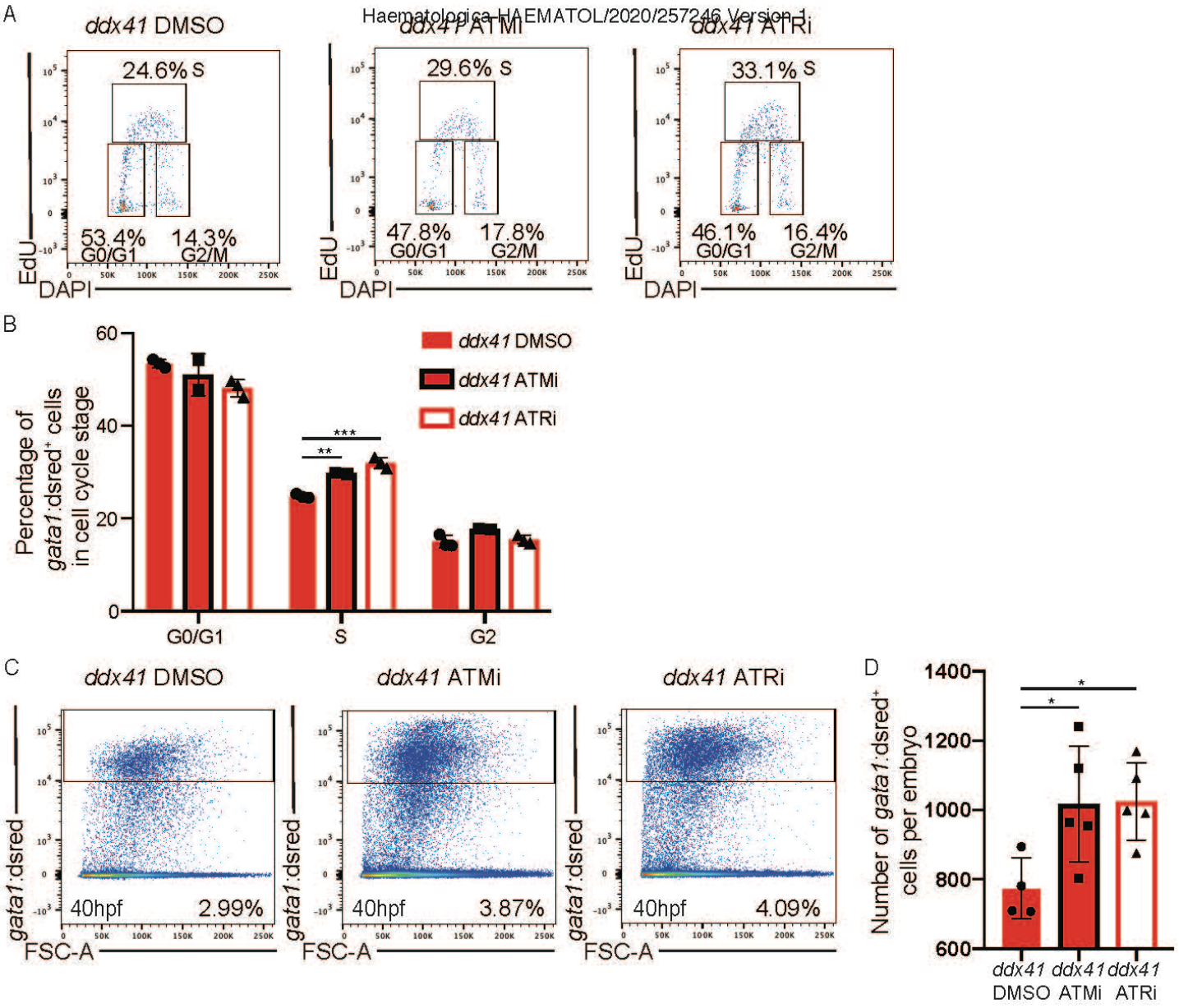
Weinreb et al. Figure 2



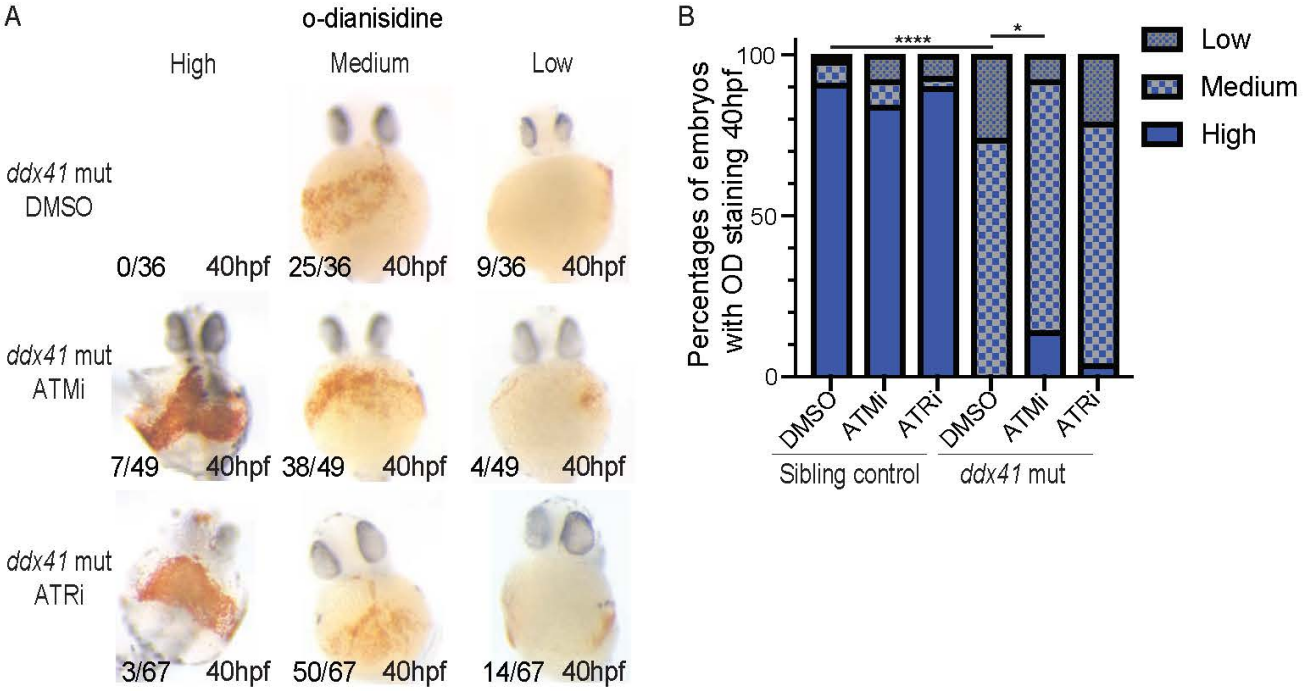
Weinreb et al. Figure 3



Weinreb et al. Figure 4



Weinreb et al. Figure 5



Weinreb et al. Figure 6